PhD-FSTM-2023-006
The Faculty of Science, Technology and Medicine

DISSERTATION

Defence held on 30/01/2023 in Esch-sur-Alzette

to obtain the degree of

DOCTEUR DE L'UNIVERSITÉ DU LUXEMBOURG

EN BIOLOGIE

by

Ahmed Abdelmonem Abdelmawla Ebrahim HEMEDAN
Born on 1.10.1989 in Kafrelshikh, Egypt.

APPLICATIONS OF BOOLEAN MODELLING TO STUDY AND STRATIFY DYNAMICS OF A COMPLEX DISEASE

Dissertation defence committee

Dr Reinhard Schneider, dissertation supervisor
Professor, Université du Luxembourg

Dr Anna Niarakis
Associate Professor at University of Evry Val d'Essonne
Biology Department INRIA delegate at the Group Lifeware

Dr Emma Schymanski, Chairman
Professor, Université du Luxembourg

Dr Enrico GLAAB
Assistant professor, Université du Luxembourg

Dr Burkhard Rost, Vice Chairman
Professor, Technische Universität München

Applications of Boolean modelling to study and stratify dynamics of a complex disease

Dissertation Defense Committee:

Committee members:	Prof. Dr. Burkhard Rost
	Prof. Dr. Emma Schymanski
	Dr. Anna Niarakis
	Dr. Enrico Glaab
	Prof. Dr. Reinhard Schneider
Supervisor:	Reinhard Schneider, full professor

Affidavit

I hereby confirm that the PhD thesis entitled "Applications of Boolean modelling to study and stratify dynamics of a complex disease" has been written independently and without any other sources than cited.

In ..... date ....

Author's signature

Acknowledgements

First of all, I would like to thank my advisor Dr. Marek Ostaszewski for his excellent guidance. Marek has a supportive personality and a positive attitude, which is essential to maintaining a successful professional relationship throughout the PhD. It won't be an overstatement to say that Marek has such a professional style to bring out the best performance in a student without becoming prescriptive in his guidance. This allows me the freedom to develop my own working style. He encouraged me to communicate in an open, honest, and fair manner about issues that arise, as they arise, and set clear expectations. When I reached a crossroads in my PhD research, he consistently helped me to achieve the most effective results.

I would like to express my sincere gratitude to my main supervisor Prof. Dr. Reinhard Schneider, and Dr. Venkata Satagopam. They provided me with great opportunities and helped me to grow as a junior researcher. They provided such a culture of recognition in our group, giving credit to those who deserve it.

I am extremely grateful to Dr. Anna Niarakis for her invaluable support, insightful suggestions and endless encouragement during my PhD journey. I also would like to thank Dr. Enrico Glaab for his great suggestions during my research. I always found him informative and responsive to my questions when I contacted him.

I would like to extend my sincere appreciation to Prof. Rudi Balling, the former director of the LCSB, for creating such multidisciplinary and stimulating environment.

I am grateful to my friends and colleagues- Miroslav Kratochvil, Vincent Gureghian, Armin Rauschenberger, Mathias Galati- the charming gardeners who make happiness and knowledge blossom.

To my parents, I would like to thank you for your unwavering support and love. Without you, I would not be the person I am today.

Above all I would like to thank my lovely wife Haneen for her love and constant support, and for keeping me sane over the past few years. Thank you for being my muse, editor, proofreader, and sounding board. But most of all, thank you for being my best friend. You've always been everything

Contents

Acronyms

- **ARFGAP1** : ADP-ribosylation factor GTPase-activating protein 1
- **ATP**: Adenosine Triphosphate
- **BF**: Boolean Function
- **BioCarta**: Database of Pathways
- **BN**: Boolean Network
- **BM**: Boolean Model
- **CD**: Cluster of Differentiation (a Protein Found on the Surface of Cells)
- **CABEAN**: Causal Boolean Network Ensemble Analyzer
- **CL**: Common Language Effect Size
- **CT**: Computerised Tomography
- **DaTSCAN**: Dopamine Transporter Single Photon Emission Computed Tomography
- **DM**: Diabetes Mellitus
- **DTW**: Dynamic Time Wrapping
- **GBA**: Glucocerebrosidase
- **GNA**: Genetic Network Analyzer
- **GTP**: Guanosine Triphosphate
- **IDH**: Isocitrate Dehydrogenase
- **IL**: Interleukin

- **iPSC**: Induced Pluripotent Stem Cell
- **KEGG**: Kyoto Encyclopedia of Genes and Genomes
- **KGDHC**:Ketoglutarate-Dehydrogenase Complex
- **LRRK2**: Leucine-rich Repeat Kinase 2
- **MAPK**: Mitogen-activated Protein Kinase
- **MaBoSS**: Markov Boolean Stochastic Simulation
- **MPBM**: Most Permissive BM
- **MPTP**: 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine
- **NADH**: Nicotinamide Adenine Dinucleotide
- **NetDS**: Network Dynamics Simulator
- **NR4A2**: Nuclear Receptor Subfamily 4 Group A Member 2
- **ODEs**: Ordinary Differential Equations
- **PAth2Models**: PAthway TO ModelS
- **PET**: Positron Emission Tomography
- **PI3K**: Phosphoinositide-3-kinase
- **PINK1**: PTEN-induced Putative Kinase 1
- **PD**: Parkinson's Disease
- **PDC**: Pyruvate Dehydrogenase Complex
- **PDKs**: Pyruvate Dehydrogenase Kinases
- **PGC-1α**: Peroxisome Proliferator-activated Receptor Gamma Coactivator 1-alpha
- **PKB**: Protein Kinase B
- **RA**: Rheumatoid Arthritis
- **REM**: Rapid Eye Movement
- **ROS**: Reactive Oxygen Species

- **rxncon**: Reaction-contingency Language
- **SAT**: Boolean Satisfiability
- **SABIO-RK**: Database of Biochemical Reactions
- **SBGN**: System Biology Graphical Notation
- **SBML**: Systems Biology Markup Language
- **SBML-qual**: Systems Biology Markup Language Qualitative
- **SCC**: Strongly Connected Components
- **SIRT3**: Sirtuin 3
- **SNCA**: Alpha-synuclein
- **SPECT**: Single-photon Emission Computerized Tomography
- **SQUAD**: Standardised QUAlitative Dynamical System Approach
- **SIF**: Simple Interaction Format
- **TF**: Transcription Factor
- **T2DM**: Type Two Diabetes Mellitus
- **Th**: T Helper Cell
- **TCA**: Tricarboxylic Acid Cycle
- **TCR**: T Cell Receptor
- **TGFB1**: Transforming Growth Factor Beta 1
- **Treg**: Regulatory T Cell
- **UBS**: Ubiquitin-proteasome System
- **Wnt/PI3k-Akt**: Wnt/phosphoinositide 3-kinase-Akt Pathway

Abstract

Interpretation of omics data is needed to form meaningful hypotheses about disease mechanisms. Pathway databases give an overview of disease-related processes, while mathematical models give qualitative and quantitative insights into their complexity. Similarly to pathway databases, mathematical models are stored and shared on dedicated platforms. Moreover, community-driven initiatives such as disease maps encode disease-specific mechanisms in both computable and diagrammatic form using dedicated tools for diagram biocuration and visualisation. To investigate the dynamic properties of complex disease mechanisms, computationally readable content can be used as a scaffold for building dynamic models in an automated fashion. The dynamic properties of a disease are extremely complex. Therefore, more research is required to better understand the complexity of molecular mechanisms, which may advance personalized medicine in the future.

In this study, Parkinson's disease (PD) is analyzed as an example of a complex disorder. PD is associated with complex genetic, environmental causes and comorbidities that need to be analysed in a systematic way to better understand the progression of different disease subtypes. Studying PD as a multifactorial disease requires deconvoluting the multiple and overlapping changes to identify the driving neurodegenerative mechanisms. Integrated systems analysis and modelling can enable us to study different aspects of a disease such as progression, diagnosis, and response to therapeutics. Therefore, more research is required to better understand the complexity of molecular mechanisms, which may advance personalized medicine in the future. Modelling such complex processes depends on the scope and it may vary depending on the nature of the process (e.g. signalling vs metabolic). Experimental design and the resulting data also influence model structure and analysis. Boolean modelling is proposed to analyse the complexity of PD mechanisms. Boolean models (BMs) are qualitative rather than quantitative and do not require detailed kinetic information such as Petri nets or Ordinary Differential equations (ODEs). Boolean modelling represents a logical formalism where available variables have binary values of one (ON) or zero (OFF), making it a plausible approach in cases where quantitative details and kinetic parameters

are not available. Boolean modelling is well validated in clinical and translational medicine research.

In this project, the PD map was translated into BMs in an automated fashion using different methods. Therefore, the complexity of disease pathways can be analysed by simulating the effect of genomic burden on omics data. In order to make sure that BMs accurately represent the biological system, validation was performed by simulating models at different scales of complexity. The behaviour of the models was compared with expected behavior based on validated biological knowledge. The TCA cycle was used as an example of a well-studied simple network. Different scales of complex signalling networks were used including the Wnt-PI3k/AKT pathway, and T-cell differentiation models. As a result, matched and mismatched behaviours were identified, allowing the models to be modified to better represent disease mechanisms. The BMs were stratified by integrating omics data from multiple disease cohorts. The miRNA datasets from the Parkinson's Progression Markers Initiative study (PPMI) were analysed. PPMI provides an important resource for the investigation of potential biomarkers and therapeutic targets for PD. Such stratification allowed studying disease heterogeneity and specific responses to molecular perturbations. The results can support research hypotheses, diagnose a condition, and maximize the benefit of a treatment. Furthermore, the challenges and limitations associated with Boolean modelling in general were discussed, as well as those specific to the current study.

Based on the results, there are different ways to improve Boolean modelling applications. Modellers can perform exploratory investigations, gathering the associated information about the model from literature and data resources. The missing details can be inferred by integrating omics data, which identifies missing components and optimises model accuracy. Accurate and computable models improve the efficiency of simulations and the resulting analysis of their controllability. In parallel, the maintenance of model repositories and the sharing of models in easily interoperable formats are also important.

Chapter 1

Introduction

Extensive amounts of omics data generated to understand disease mechanisms require interpretation to formulate meaningful hypotheses[1]. Pathway databases [2, 3, 4] give an overview of disease-related processes, while mathematical models give qualitative and quantitative insights into their complexity. Similarly to pathway databases, mathematical models are stored and shared on dedicated platforms [5, 6, 7, 8, 9]. Moreover, community-driven initiatives such as disease maps [10] encode disease-specific mechanisms in both computable and diagrammatic form using dedicated tools for diagram biocuration [11, 12, 13] and visualisation[14, 15]. In all cases, computationally readable content can be used as a scaffold to build dynamic models in an automated fashion to investigate the dynamic properties of the system [16]. The dynamic properties of a disease are extremely complex. Therefore, more research is required to better understand the complexity of molecular mechanisms, which may advance personalized medicine in the future. Modelling of such complex process depends on the scope and it may vary depending on the nature of the process (e.g.signalling vs metabolic). The experimental design and resulting data also influence the model structure and analysis [17]. Dynamic modelling approaches include Boolean or Multi-valued models [18], Petri nets [19] or Ordinary Differential equations (ODEs)[20]. However, model parameterisation is a challenging task [21] making logical models an interesting alternative[8].

1.1 The complexity of Parkinson's disease

As an example of complex disorders is Parkinson's disease (PD). PD is characterized by progressive damage to dopaminergic neurons, causing motor, cognitive, and behavioral impairments. PD is characterized by a loss of neurons in the substantia nigra, resulting in rigid muscles, tremors, and bradykinesia. The de-

generation of dopaminergic neurons in the substantia nigra over time decreases dopamine release, resulting in reduced stimulation of the motor cortex by the basal ganglia[22].

Although the causes of PD are unclear, evidence suggests that genetic and environmental factors contribute to its development. The majority of PD cases are sporadic rather than familial, and are caused by environmental factors or a combination of several genetic and environmental factors. Genetic mutations are found to be responsible for up to 15 percent of PD cases[23]. For example, LRRK2 mutations increase the aggregation of cytosolic proteins, leading to apoptosis and cell dysfunction [23].

The cause of PD is not only genetic, but is also linked to lifestyle factors and toxin exposure. As an example, MPTP (1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine) is an environmentally-derived neurotoxin [23]. MPTP destroys dopaminergic neurons, leading to permanent symptoms similar to PD. There is no definitive link between environmental toxins and PD progression, but it is hypothesized that exposure to environmental toxins increases the risk[23].

A variety of comorbidities are associated with PD, including cardiovascular diseases, melanoma, dementia, and diabetes [24]. The most common comorbidities among PD patients are hypertension and diabetes [24]. In some cases, depression may appear decades before the onset of PD. Limited information is known about how comorbidities may play a role in PD pathogenesis and progression. PD and its comorbidities may share therapeutic targets and biomarkers. Diabetes-related drugs are currently being evaluated in clinical trials for PD [25].

From the previous examples, PD is associated with complex genetic, environmental causes and comorbidities that need to be analysed in a systematic way to unveil the progression of different disease subtypes. Studying PD as a complex disease requires deconvoluting the multiple and overlapping changes to identify the driving neurodegenerative mechanisms. This requires multiple and precise perturbation experiments to study the effect of mutations on PD mechanisms, which is not possible in the human body[26]. For this reason, scientists include model systems, such as in vivo and culture systems, to identify novel mechanisms that will lead to new therapeutic approaches. In this context, multi-dimensional omics analysis of culture experiments is essential to provide a large number of measurements for relatively few samples. Further, integrated systems analysis can help us to study many aspects of a disease in parallel. Systems analysis of human tissues yields a variety of novel insights regarding PD-related mutations.

Analysis of other invivo models yields correlations between measurements and disease phenotypes. Therefore, a multivariate analysis is necessary to identify many correlates of disease and discriminate between disease subtypes [27]. This leads to decreased motor control and movement. As PD progresses, cognitive impairment, depression, and other behavioral symptoms can develop.

1.2 Logical modelling process

Logical modelling in systems biology is a mathematical approach that can be used to represent biological systems. It is based on Boolean algebra and can be used to represent the interactions between components in a system. This approach is useful for understanding how components interact and how changes in one component can affect the whole system. Logical models are often used to investigate the behaviour of cellular networks or other systems, such as gene regulation networks. Logical models are useful for predicting the effects of perturbations to a system, such as drug treatments, and can provide insights into the dynamics of a biological system.

There are several types of logical modelling used in systems biology. Boolean models (BMs) are the most commonly used type of logical modelling. These models provide a graphical representation of a system and are used to model the interactions between components. Probabilistic BMs are an extension of BMs and allow for the incorporation of probabilistic functions. Petri nets are another type of logical modelling used in systems biology. These models provide a graphical representation of a system and are used to model the interactions between components and the effects of different inputs. Finally, rule-based models are used to represent the dynamics of a system and the behaviour of components in response to external stimuli.

BMs are qualitative rather than quantitative and do not require detailed kinetic information such as Petri nets [19] or Ordinary Differential equations (ODEs)[20]. However, in some research areas, such as pharmacogenomics, presenting data to simple BMs may be challenging, and does not introduce the best description of the biological system [28]. Therefore, researchers studied the qualitative nature of BMs, facilitating the integration with other quantitative methods to allow better analysis[29]. Such methods, including ODEs and Petri nets, combined with BMs and constraint-based models, show that BMs are useful scaffolds for quantitative models [28].

Boolean modelling represents a logical formalism where available variables have binary values one (ON) or zero (OFF), making it a plausible approach in

cases where quantitative details and kinetic parameters are not available. Model's biomolecules and interactions are described by Boolean functions (BFs) that define the behaviour of the output biomolecules based on the interaction inputs. Some non-physical components of BMs, represented as pathway endpoints or phenotype, are not biomolecules. Updating schemes define conditions and order in which the BFs are calculated [30]. BFs were used primarily to describe genes' regulation but other researchers applied them to signalling networks using various logical formalisms [31, 32, 33], e.g. Boolean, differential, or fuzzy equations.

Boolean modelling was applied in clinical and translational medicine research[34, 35, 36]for various purposes. Simulation of the complex biological systems allowed to predict the activity of pathway endpoints (phenotypes)[37], drug targets [38] and cellular crosstalks [39]. Identifying attractors helped to understand the activity of the phenotypes, since they represent the steady states of biomolecules [40, 41]. Finally, comparing attractors before and after perturbations allows evaluating the model stability and give insight into how the in-vivo systems maintain their homeostasis.

In this project, the PD map is translated into BMs in an automated fashion using different methods. Therefore, the complexity of disease pathways is studied by simulating the effect of genomic burden from omics data. BMs are created at different scales of the complexities to ensure its ability and reliability to simulate disease mechanisms. First, the simple and known mechanisms are simulated to investigate the model ability to represent the already known behaviour. Further, the BMs are used to re-simulate complex molecular interventions data, comparing the results with other models.

The BMs are stratified by integrating omics data of multiple disease cohorts. Such stratification allows for studying disease heterogeneity and specific responses to molecular perturbations. Molecular expressions can be used to specify the conditions of key components such as disease biomarkers. The effect of presence or absence of disease biomarkers is simulated, modifying the probability of having a component active or inactive at the beginning of the simulations. From this point, the modelling of disease modules can develop specific treatment strategies and propose new therapeutic interventions to arrest disease progression.

Taken together, Boolean modelling is a flexible formalism, allowing to analyse a range of dynamic properties of biological systems, including models of disease mechanisms[42, 43]. BMs show a flexibility to be integrated with omics data, allowing a precision in transitional medicine.

1.3 Result overview

BMs were created in different modelling formats including SBML-qual, which is used for creating, storing and exchanging qualitative models. Analysis of the models' structural and dynamic properties was used to verify their accuracy. The dynamic verification and sensitivity analysis results showed that the BFs accurately represented the original interactions and the model was robust against small perturbations. The validation was performed by simulating models in different scales of complexity. The behaviour of the models was compared with the expected behavior based on validated biological knowledge.

The BMs were stratified based on the dataset obtained from the Parkinson's Progression Markers Initiative study (PPMI). PPMI provides an important resource for the investigation of potential biomarkers and therapeutic targets for PD. High-throughput sequencing of small RNAs from whole blood samples of patients at various stages of the disease was performed, and miRNAs were identified and quantified. Differential expression analysis revealed the dysregulation of several miRNAs in the blood of PD patients compared to healthy controls. In addition, the expression of several miRNAs was associated with clinical parameters of disease subtypes.

The simulation of BMs highlights different regulatory modules during disease progression. The results may support the research hypothesis, diagnosis, and maximize the effectiveness of a treatment. In addition, the models provided stage-based simulations of disease progression and diagnosis. Furthermore, the study identify the similar probabilities between phenotypes across disease subtypes. In this way, it is possible to use the models to investigate similarity-based differential diagnosis and identify common crosstalk. The approach described in this thesis is illustrated in (Figure 1.1). This figure summarizes the key steps and methods used in our research

As will be seen in the following chapters, the literature review serves as an essential foundation for the rest of the study, as it provides a comprehensive overview of the current state of knowledge on the Boolean modelling. I systematically reviewed and analysed the relevant literature, highlighting key findings and theories that are relevant to the research question. This includes a critical evaluation of the strengths and limitations of the previous research, and to synthesise the findings in a way that sets the stage for the research question and methods of the current study.

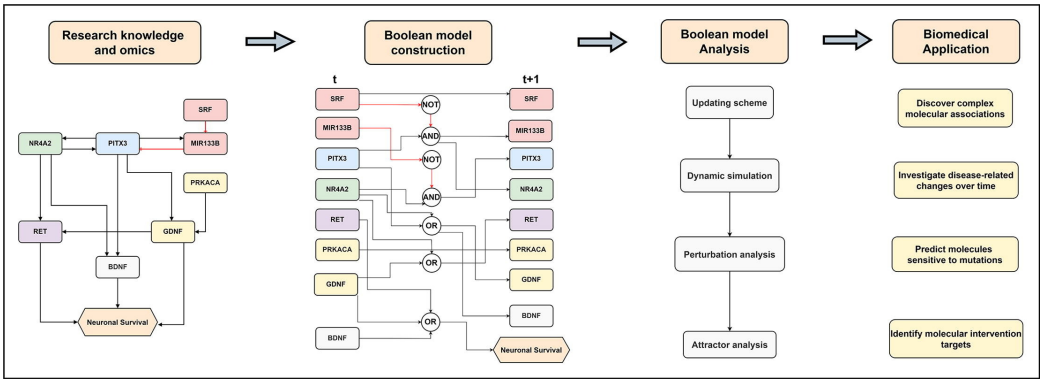


Figure 1.1 The figure summarizes the general approach taken in this thesis. It highlights the key steps and methods used in our research. The static diagrams represented as research knowledge and omics can be converted into dynamic Boolean models, enabling in silico simulations and predictions to be made. This allows for a deeper understanding of the underlying mechanisms of health and disease.

The method chapter provides the details of how the study is conducted. the study design is described, including the research question and the hypotheses being tested. I describe the datasets that were used, as well as the data collection methods and any data analysis techniques that are applied. It is important to note that the scripts that used in the method chapter are provided on Gitlab repository so that other researchers can understand and replicate the study if needed.

The results chapter include tables and figures to help illustrate the results. The chapter highlights the limitations and potential sources of error in the analysis process, as well as any unexpected or surprising results.

The discussion chapter delves into the dynamics and interpretation of the results, offering insight into the underlying mechanisms or trends that are observed in the study. This includes a discussion of the implications of the results for the research question or hypotheses being tested, as well as how the results compare to previous research in the field. The chapter explains the limitations or challenges that were encountered in the study and how they may affect the results.

The conclusion chapter highlights the key takeaways from the research, including the limitations and future directions for study.

Chapter 2

Literature review

2.1 Parkinson's disease as a complex disorder

Parkinson's disease (PD) is a complex progressive neurological disorder that affects movement and coordination. The cause of PD is not known, but it is believed to be related to a combination of genetic and environmental factors. It is characterized by the slow degeneration of nerve cells in the brain that produce dopamine, a neurotransmitter that helps control movement. As these cells are lost, the patient experiences a variety of symptoms, including tremors, slow movement, stiffness, difficulty with balance and coordination, and problems with speech [44]. The complexity of PD makes it difficult to diagnose, as symptoms can vary greatly from person to person. Diagnosis often requires a combination of tests, such as MRI, CT scans, and neurological exams [45]. Both scans cannot definitively diagnose PD, but they can assist physicians in eliminating other potential causes of the symptoms.Treatment for PD includes medications to increase dopamine levels, physical therapy to help maintain muscle strength and coordination, and lifestyle changes to manage symptoms. The complexity of PD also makes it difficult to predict how it will progress over time. While some patients may experience only mild symptoms and a relatively slow progression of the disease, others may experience more severe symptoms that worsen quickly. This unpredictability can make it difficult to plan for the future and make decisions.

2.1.1 Hallmarks and genetic mutations

PD-related pathways involve a wide range of molecular changes. These pathways are linked to the degeneration of dopamine-producing neurons in the substantia nigra in the midbrain. Research into these pathways is ongoing, and new treatments are being developed to address the underlying causes of this disorder.

Mitochondrial dysfunction

Mitochondrial dysfunction is implicated as a major contributor to the pathogenesis of PD. Mitochondrial dysfunction refers to a decrease in the ability of mitochondria to produce ATP, the primary source of energy in the body [46]. This can lead to impaired function of cells, including dopamine-producing neurons, which may contribute to the development of PD. Mitochondrial dysfunction can be caused by a variety of factors, including genetic mutations, environmental toxins, and oxidative stress [47]. Treatments aimed at addressing mitochondrial dysfunction in PD may require the use of dietary supplements and antioxidants, as well as drugs that target specific mitochondrial pathways. In addition, research show that targeting the mitochondria may be a potential therapeutic strategy for PD [47]. For example, research suggests that increasing levels of the PGC-1α protein, which plays a role in mitochondrial biogenesis and energy production, may be beneficial for PD [48]. Additionally, studies suggests that drugs that can target mitochondrial function, such as the antioxidant CoQ10, may be beneficial for PD [49, 50]. Finally, research shows that modulating the energy metabolism of mitochondria may be a promising strategy for treating PD [51].

Oxidative stress

The accumulation of reactive oxygen species (ROS) leads to the formation of toxic compounds that can damage neurons, leading to the development of PD. Oxidative stress is thought to be a major contributor to the pathology of PD. The imbalance of ROS and antioxidants leads to increased cellular damage and dysfunction of dopaminergic neurons. This increased oxidative stress is thought to contribute to the death of these neurons, leading to the clinical features of the disease. There is evidence that antioxidant supplements, such as vitamin E and vitamin C, can reduce the oxidative stress associated with PD. This may be a potential supportive treatment for the disease. Additionally, dietary interventions, such as consuming foods rich in antioxidants, may provide some protection from the oxidative damage seen in PD. Therefore, it is important to understand the role of oxidative stress in the development of PD. In addition, it is important to explore strategies to reduce the oxidative damage associated with the disease.

Neuroinflammation

Neuroinflammation is believed to be an important factor in the pathogenesis of PD. The presence of chronic inflammation in the brain is linked to the progression of the disease and the development of motor symptoms. Studies show that individuals with PD have increased levels of inflammatory mediators in their brains[52]. Additionally, genetic mutations associated with PD are linked to increased levels

of pro-inflammatory cytokines [52]. Furthermore, environmental factors such as pesticides, heavy metals and other toxins may trigger an inflammatory response in the brain which could cause the development of PD [53].

Protein aggregation

Protein aggregation plays a role in the pathogenesis of PD. The formation of protein aggregates, such as Lewy bodies, is linked to the progression of the disease and the development of motor symptoms. Protein aggregation occurs when proteins misfold and clump together, leading to the formation of insoluble structures within the cell [54]. The accumulation of these aggregates disrupts normal cell functioning, leading to cell death and the development of Parkinson's symptoms. Additionally, the presence of these aggregates contributes to a cascade of oxidative stress, inflammation and mitochondrial dysfunction, which can further aggravate the disease [55]. In addition to Lewy bodies, other aggregated proteins associated with Parkinson's include alpha-synuclein, tau and polyglutamine [54]. Research is ongoing to understand how these aggregates form, as well as to develop treatments that target them.

Protein degradation

Protein degradation is an important process in the body that helps to clear out damaged or unnecessary proteins. In PD, two types of protein degradation pathways are thought to play a role in the degeneration of neurons: ubiquitin-proteasome system (UBS) and autophagy [56]. The UBS is a pathway that involves the attachment of a small protein called ubiquitin to target proteins, which marks them for degradation by the proteasome [57]. The proteasome is a complex of enzymes that breaks down the target proteins into smaller peptides, which can then be recycled or used for energy. In PD, the UBS is thought to play a role in the degeneration of dopamine-releasing neurons, which are responsible for producing the neurotransmitter dopamine [57]. Autophagy is another pathway that plays a role in protein degradation and is important for maintaining cellular homeostasis. Autophagy involves the formation of vesicles called autophagosomes, which engulf and recycle damaged or unnecessary proteins and organelles [58]. In PD, autophagy is thought to be impaired, which may contribute to the accumulation of abnormal proteins such as alpha-synuclein. Both the UBS and autophagy pathways are important for maintaining protein homeostasis in the body and may play a role in the development and progression of PD. Further research is needed to fully understand the role of these pathways in PD and to identify potential therapeutic targets for the treatment of this condition.

Genetic components of Parkinson's disease

Research suggests that genetic factors play an important role in the development of PD. Mutations in certain genes are associated with an increased risk of developing the disease. These mutations can affect the proteins involved in dopamine production, as well as other proteins involved in the pathways outlined above. For example, mutations in the LRRK2 gene are associated with late-onset PD, which develops after age 50 [59]. LRRK2 mutations increase the aggregation of cytosolic proteins, leading to apoptosis and cell dysfunction [23].

Parkin, which is encoded by the PARK2 gene, is associated with the degradation of dysfunctional mitochondria. There is evidence that mutations in the PARK2 gene are associated with an earlier onset of PD [60]. Researchers suggest that the loss of parkin activity interferes with the normal functions of cells, including the activity of mitochondria. Damaged mitochondria may affect dopamine transmission in dopaminergic neurons, resulting in PD symptoms and signs [60].

The SNCA gene encodes the alpha-synuclein protein, which is found in presynaptic nerves and other types of cells. Alpha-synuclein plays a key role in neurotransmission as it regulates synaptic vesicle release - the release of key neurotransmitters, such as dopamine [23, 60]. SNCA mutations can lead to abnormal accumulation of dopamine due to the dysregulation of the alpha-synuclein protein. The body then breaks down what is thought to be excessive dopamine, resulting in neuronal cell death and the characteristic signs and symptoms of PD [23, 60].

The PARK7 gene codes for a protein that acts as an antioxidant, protecting neurons from oxidative stress and preventing alpha-synuclein accumulation. Mutations in the PARK7 gene lead to the accumulation of alpha-synuclein and the degradation of excess dopamine [23, 60]. Dysfunction of the PARK7 gene leads to oxidative stress, causing dopaminergic neuronal death [60].

PINK1 gene mutations may play a role in causing early-onset PD. The PINK1 gene codes for a protein, PINK1, which is found in the mitochondria of cells throughout the body. PINK1 plays a protective role in response to oxidative stress [23, 60]. Normal PINK1 protein inhibits apoptosis. However, a mutant PINK1 protein may contribute to increased neuronal cell death.

2.1.2 Environmental factors related to Parkinson's disease

Environmental factors, such as exposure to certain chemicals and toxins have are linked to an increased risk of PD. Other environmental factors may also play a role, such as lifestyle factors, like exposure to certain pesticides. MPTP is an environmental neurotoxin that is linked to Parkinson's-like symptoms in humans and animals. It inhibits the enzyme responsible for breaking down dopamine.

This leads to a buildup of dopamine in the brain, which causes cell death and the characteristic symptoms of Parkinson's. In addition to its potential link to the development of PD. It can also be used to study the effects of dopamine-related treatments in animals. Despite its potential link to PD, further research is needed to confirm any causative relationship [23].

2.1.3 Comorbidities associated with Parkinson's disease

PD can result in a variety of comorbidities, or additional health conditions that can occur alongside the primary diagnosis. Common comorbidities in people with PD include depression, anxiety, urinary incontinence, sleep disturbances, dementia, and REM behavior disorder. Treating these comorbidities can help improve the quality of life for people with PD and reduce the risk of complications [24].

Although there is limited research on the comorbidity of PD and Type 2 Diabetes Mellitus (T2DM), there is evidence that suggests that the two conditions co-occur more often than is expected by chance alone [61]. There are a few potential explanations for the comorbidity of both conditions. Firstly, the mechanisms underlying both conditions are similar and linked. For example, certain pathways involved in the development of PD such as oxidative stress, mitochondrial dysfunction, and inflammation, are associated with the pathogenesis of T2DM. Additionally, certain risk factors for both conditions overlap, such as age and genetics [62, 63]. Lastly, certain drugs used to treat PD, such as levodopa and dopamine agonists, may increase the risk of T2DM [64].

T2DM is a metabolic condition in which the body does not produce enough insulin to effectively regulate blood sugar levels. T2DM may worsen the motor symptoms of PD [63]. This can include a decrease in mobility, balance, and coordination. It can also lead to an increased risk of cognitive decline [65]. T2DM has a significant negative impact on the lives of people with PD. To reduce the risk of these complications, PD patients should be checked for signs of T2DM and treated appropriately. Additionally, lifestyle modifications such as a healthy diet and regular exercise should be encouraged to help control blood sugar levels.

PD is frequently associated with insulin resistance [63]. Insulin resistance is a condition in which the cells become less sensitive to the effects of insulin. This can lead to high blood sugar levels, which can increase the risk of T2DM and other health complications. Studies show that PD patients may have increased insulin resistance due to a decrease in insulin secretion, changes in fat metabolism, increased oxidative stress, and increased inflammation [66, 67]. Additionally, PD patients may have an increased risk of developing T2DM due to lifestyle factors such as physical inactivity and a poor diet. Treatment of insulin resistance in PD patients can involve medications such as metformin and insulin sensitizers can

be used to help lower blood sugar levels and improve insulin sensitivity [68].

2.1.4 Cohort study of Parkinson's disease

From the previous examples, PD is associated with complex genetic, environmental factors, and comorbidities. These complexities need to be analysed in a systematic way to understand the progression of different disease subtypes. Cohort studies are an effective tool for understanding the multiple interactions of complex diseases. These studies can help identify risk factors associated with the development of a particular disease. They can also be used to assess the effects of interventions designed to reduce the risk or progression of the disease. Cohort studies are especially helpful in complex diseases because they allow researchers to observe the progression of the disease over time. In addition, they help to identify any changes in risk factors that may be associated with the onset or progression of the disease. Cohort studies can also provide valuable insight into the biological mechanisms underlying the disease, and can be used to develop more effective treatments or preventative strategies.

In this context, the Parkinson's Progression Markers Initiative (PPMI) is an observational, international, longitudinal study of PD. The goal of PPMI is to identify biomarkers that could predict the onset of PD and measure the progression of the disease. To achieve this goal, PPMI collects data from participants on a variety of measures, including demographic information, clinical assessments, laboratory tests, and imaging studies. PPMI also collects biological samples, such as blood, urine, and saliva, which are analyzed to identify biochemical markers of PD [69]. The data and biological samples obtained at PPMI are made available to researchers around the world through the PPMI Data and Sample Repository (https://www.ppmi-info.org/access-data-specimens/data). This repository provides access to data sets, including clinical, demographic, imaging, and laboratory information, as well as protocols, study materials, and biospecimens. The repository also provides access to analytical tools, such as data visualization and bioinformatics software, to aid in biomarker discovery.

As part of PPMI, blood samples are collected from participants and analyzed to measure the levels of microRNAs, which are small and stable molecules that regulate gene expression. These microRNA levels can be used to determine disease subtypes as well as to monitor the progression of the disease [70]. The PPMI includes different disease subtypes such as SWEDD (scans without evidence of dopaminergic deficit) and prodromal cases (early stage of a disease).

SWEDD (187 individuals) refers to the absence, rather than the presence, of an imaging abnormality in patients clinically presumed to have PD. SWEDD is a neurologic diagnosis made in patients who have a classic clinical presentation of PD but lack the classic imaging features seen in PD on a brain scan [71]. This can occur when a patient has symptoms of PD but does not have the usual brain imaging findings seen in PD. For example, the presence of Lewy bodies and/or the presence of a midbrain dopaminergic deficit on PET or SPECT scans.

Prodromal (223 individuals) does not usually develop severe symptoms, but it can be diagnosed through a positive dopamine transporter (DAT) SPECT scan. A DAT SPECT scan is a type of imaging test that measures the activity of dopamine transporters in the brain. This test can help diagnose prodromal, as it can detect changes in dopamine transporter activity that may indicate prodromal. It is important to note that a positive DAT SPECT scan does not necessarily mean that a person has prodromal symptoms. It means that the scan detected changes in dopamine transporters which may indicate a prodromal condition [72].

Other cohort studies includes the NCER-PD cohort[73] and OPDC Discovery cohort [74]. NCER-PD is known as the Luxembourg Parkinson's study, aims to conduct comprehensive clinical, molecular, and device-based research on PD and atypical parkinsonism [73]. This study includes patients of all stages of the disease, regardless of age, comorbidities, or linguistic background. To facilitate this research, the NCER-PD cohort implements an open-source digital platform that can be integrated with other international PD cohort studies. Additionally, the NCER-PD cohort is focusing on specific areas of PD research relevant to Luxembourg, including vision, gait, and cognition [73]. The NCER-PD cohort also collects high-quality biosamples, such as body fluids and tissue biopsies, to support its research efforts. The combination of advanced clinical phenotyping approaches and device-based assessment allows for a comprehensive understanding of the disease and its progression. The OPDC Discovery cohort is a long-term study that is researching PD. The study includes individuals with early-stage Parkinson's, people at risk of developing the condition, and those with REM Sleep Behaviour Disorder, as well as healthy controls [74]. The goal of the study is to learn more about how Parkinson's progresses in different individuals and to use this information to improve diagnosis and management of the disease. The OPDC is committed to using new and innovative technologies to aid in the diagnosis and monitoring of Parkinson's symptoms. The hope is that these technologies will help affected individuals better manage their symptoms and have more control over their disease.

2.2 Logical modelling in systems biology

Modelling biological processes can take various forms depending on the scope of the study and the questions being explored. Data and prior knowledge of pathways also play a role in the structure and analysis of the model[75]. Static network analysis, for instance, is used to identify clusters and correlations between biomolecules. This approach can be useful for understanding the relationship between functional biomolecules and disease classes or phenotypic measurements [76]. Dynamic modelling focuses on simulating the causal relationships between biomolecules and their measurement values. By integrating data and prior knowledge of pathways, dynamic models can be used to model relationships between expressed genes or proteins [76]. This approach can be applied to understand the effects of a particular drug on the progression of a medical condition. In this case, data such as drug concentrations and biomarkers can be used to construct and modify the model, allowing it to be used as a tool to explore hypotheses and identify potential therapeutic targets [77].

There are different types of dynamic modelling approaches, for example, Petri nets [78] or Ordinary Differential equations (ODEs) [79] and Boolean, Multi-valued models [80]. However, model parameterisation is a challenging task [81] making Boolean or multi-valued discrete models, an interesting alternative [82]. Petri nets are a mathematical modelling approach that is frequently used in systems biology to understand and analyze complex biological systems [83]. Essentially, a Petri net consists of a directed bipartite graph with nodes representing different biomolecules (such as proteins, DNA, or RNA) and transitions representing the interactions or events that can take place within the system. Edges in the graph represent the flow between biomolecules. By simulating the flow through the graph, scientists can study the dynamic behavior of the system and make predictions about how it will behave under different conditions [78]. Petri nets are applied to a wide range of biological processes, including gene regulation, protein synthesis, and cell signaling pathways, and are useful for identifying potential targets for therapeutic intervention [84].

ODEs are type of mathematical equations that used to model the dynamics of biological systems, such as the concentration of biomolecules [79]. For example, an ODE may be used to describe the change in the concentration of a particular chemical compound in a pathway over time, taking into account the rate at which it is produced, consumed, and eliminated [21]. ODEs can also be used to identify key factors that influence the behavior of a biological system, and to design experiments or interventions that can alter the system's behavior in desired ways [85].

A Boolean model (BM) does not require detailed kinetic information since it is qualitative rather than quantitative. It can be challenging to present unknown data to simple BMs in certain research areas, such as pharmacogenomics [28]. As a result, researchers explored how BMs can be integrated with other quantitative methods to make analysis more effective [86, 87].

Biomolecular systems can be represented by BMs, where biomolecules' states have two possible values, one or zero, that change according to their interactions described by Boolean functions (BFs). A BF defines the output state based on inputs and the logic of the interaction. The order of evaluating BFs is governed by an updating scheme [29]. BMs are used to analyze qualitative behaviour and therapeutic responses to a variety of biological problems [31, 33, 32]. Here, we review the application of Boolean modelling to systems medicine problems. An overview of the modelling process is followed by an explanation of incremental development of the Boolean modelling in clinical and translational medicine. We conclude by discussing emerging tools and methods for improving the reproducibility and reuse of such models in biomedical research.

2.2.1 Building the regulatory graph of the Boolean models

In general, the first step of Boolean modelling is building model structure. This step can be achieved by inference of molecular interactions from available mechanistic data [30, 88] The second step is to construct BFs that describe the interactions of biomolecules. The transition of a biomolecule from one state to another is regulated by those functions and the chosen updating scheme.

Inference from molecular interaction diagrams

Molecular mechanisms of health and disease are often represented as systems biology diagrams, and the coverage of such representation constantly increases. These diagrams can be transformed into BMs using automated conversion tools. Simple interaction format (SIF) can be translated into a list of logical functions in Boolsim format [89]using a Standardised QUAlitative Dynamical system approach (SQUAD) [90]. In turn, diagrams and networks from KEGG, BioCarta and SABIO-RK [91] can be transformed to SBML-qual format using PAth2Models, providing large scale models. SBML-qual offers the interoperability and the annotation of a BM and its components. The System Biology Graphical Notation (SBGN) Activity Flow, Process Description and entity relation diagrams can be translated to the reaction-contingency language [92] (rxncon language), using the mechanistic details on the signal transduction networks and creates the contingencies (logical rules).

Recently, the diagrams in CellDesigner SBML format were translated to SBML-qual using CaSQ (CellDesigner as SBML-qual), providing the construction of large-scale BMs automatically from the molecular interaction diagrams. The translation process starts with the diagram reduction based on specific rewriting rules [93]. CaSQ also infers the Boolean functions and translates the interactions from the Process Description to the Activity Flow, considering one qualitative species to each species after the reduction step.

Construction of Boolean functions

In a molecular setting, biomolecules are controlled by various regulators, which can be represented as regulatory nodes controlling the activation or inhibition of their target nodes by associated rules called BFs. The state of a specific biomolecule is changed based on BFs in an iterative manner [94]. Each iteration corresponds to a discrete time step in a path length (the number of the interactions in a path)and completes as long as all biomolecules are updated based on a proper time scheme [30].

Other types of BFs that are not as widely used in systems biology, as they describe complex and non intuitive relationships, are described in detail in Shuster (Ed.) [95]. An important type among these is a canalysing function [96]. In this type, multiple input variables that control the BF are in a hierarchical relationship [96]. For example, in A OR (B AND C), if A equals one, it will control the output regardless of other variables. If A equals zero, the second part of the function will be controlled by B= 0 or C=0. It was noticed that the canalysing functions increase robustness of BMs, similarly to corresponding molecular mechanisms. BFs can also be constructed by different probability distributions [95]. They can represent the combinatorial effects of gene regulations in a simple and interpretable representation of regulatory networks [97].

2.2.2 Simulation of model dynamics with updating schemes

The dynamics of BMs are simulated by incremental execution of BFs and changing the states of corresponding nodes in a series of discrete time steps. The order at which BFs are executed is governed by updating schemes: synchronous, asynchronous and hybrid [98]. Updating schemes are fundamental in dynamic models to understand the mechanism of how the biomolecules are evolved in a biological system across time [99]. The evolution of one biomolecule may not be described by only one interaction, therefore, the modeller can test and choose different options of updating schemes to decide which scheme is more suitable for the research.

The synchronous scheme updates the state of all biomolecules at the same

time. The resulting dynamic trajectory of a synchronous updating scheme is deterministic i.e. the model always reaches the same state after the same time.

The asynchronous scheme updates the state of one biomolecule per transition, and the nodes are updated randomly based on their BFs. Therefore, one initial condition state can evolve into different outcomes with different executions [100]. The spectrum of all possible outcomes is illustrated using the state transition graph.

In hybrid scheme, synchronous and asynchronous schemes can be used together. It includes temporal logic to model interactions that need more than one time step [101, 102]. The temporal logic adds time delays between regulator biomolecules and their final products. The hybrid scheme assumes partitioning the model into clusters, the variables at each cluster are synchronously updated and the choice to start the update between groups relies on asynchronous strategy [101, 103].

The previous schemes can be updated in probabilistic manner using probabilistic BMs. Probabilistic BMs assigns certain probabilities to BFs, and each biomolecule is updated based on this probability before reaching the steady state [101, 104].

2.2.3 Attractor analysis

A simulated model can reach a stable dynamic behaviour, where the states of the biomolecules converge to a stable configuration, called an attractor, which is interpreted as a physiological endpoint [105, 34]. An attractor is a state of a BM with no outgoing edges in the state transition graph. Attractors can be classified as i) stable states (fixed points) which are time invariant, and ii) complex attractors – sets of possible outcomes that can be reached following the synchronous and asynchronous scheme [106]. The set of states within an attractor is called the basin of attraction. It can be interpreted as a set of possible biological scenarios, supporting testable hypotheses [107]. In synchronous and deterministic asynchronous schemes, the system may oscillate regularly when attractors form a limit cycle, and each node has not more than one successor. An example of a limit cycle is the cell cycle in models of a eukaryotic cell [108, 109]. In a stochastic asynchronous scheme, the system may oscillate irregularly due to the random initial condition selection leading to loose attractors. That means the model does not oscillate in a cycle due to the target node having more than one successor. It is challenging to interpret complex attractors with large numbers of steady states that oscillate in an irregular cycle.

To find an attractor, the past states of the model are compared to the updated ones to find recurring patterns. This search process can be exhaustive or heuristic. An exhaustive search starts from all states synchronously until the attractor is

reached. This mode is mostly limited to small-size models [102], although a SAT solver can increase the search speed, identifying the possible attractors in large models with hundreds of components [110]. In turn, the heuristic search starts with a chosen subset of states to identify the attractor synchronously or asynchronously. The heuristic search performs random transitions, creating model states with a high probability. Then, the algorithm computes the forward reachable sets of the model states. If all sets are similar, an attractor is identified [111].

Identifying an attractor in a complex model is challenging. Many reduction techniques were implemented to simplify the original BFs to include a fewer number of operations [112, 113, 114]. This can be achieved by removing components that do not affect the behaviour of the original BFs. For example, some reduction techniques identify the biomolecules whose state do not change, turning the corresponding BFs into a simpler model [112, 115]. In complex BMs, this technique is followed by removing interactions with one input and output and self-loops [116]. Another approach splits the model into strongly connected components (SCCs) to decrease its complexity, and the simulations are run for all the SCCs independently [117]. Recently proposed Most Permissive BMs simulations is a paradigm to perform trajectories sampling and to reach the complete set of attractors faster than the asynchronous search, allowing to run more fine-grained simulations [118].

2.2.4 Topology, perturbation, and controllability analysis

Simulation of BMs to identify their stable states and attractors provides insights into the behaviour of the model. From this point, it is possible to predict meaningful interventions towards desired outcomes by analysing the structure of the model and its response to perturbations. To gain insight into the model structure, it is important to study the topology of a BM which may be a necessary prerequisite for some updating schemes and/or attractor analysis (the ones that need modularisation). Such analysis helps to understand the connectivity of the model components and how they affect phenotypes. This information can improve the understanding of BM dynamics under different updating schemes or attractor analysis [119] by identifying structural cycles in the BM topology. Moreover, this information can be used to define components sensitive against perturbations [120].

Perturbation analysis means changing the state of a biomolecule or its BFs, to analyse the topological robustness and the dynamic resilience of the BM, and the attractors it reaches [121, 122]. Comparing the original attractors with those after perturbation allows evaluating its impact. One of frequently used perturbations sets the state of a biomolecule to a fixed value, zero or one, emulating permanent

activation or activation, e.g., due to a drug action. Other types of perturbation may change the rule structure of the BFs, either entirely (rule-flip) or partially. Such partial perturbations are called edge perturbations, as they affect the connectivity of a BM.

Testing of BMs is performed at three levels. The first level is to check the model topology of the biomolecules and their BFs for completeness and correct use of logic operators [123]. The second level is to ensure the model ability to simulate a biological system at different updating schemes. The third level is to check the stability of a model in response to perturbations. A robust model reflects how the physiological system can keep their internal state relatively constant against the effect of perturbations [30]and maintain homeostasis.

The control of BMs can be achieved by two main ways. The first one starts with adding external sets of signals to affect the state of the biomolecules in the model and reach the desirable stable state or attractor [124, 125]. The added signals, represented as additional nodes in BMs, have no parent interactions and their values are a series of state values in a specific time step to reach the desirable states. They can represent possible therapies affecting model behaviour. For instance, the control of gene expression is essential to plan therapeutic interventions [124].

The second way is to perturb the states of BMs randomly to select the biomolecules that may result in attractors representing the desired outcomes of the model. This approach is implemented as an algorithm [126]that that identifies the optimal one-bit perturbation, i.e., the simplest form of perturbation that inverts the states of biomolecules in an attractor, for a given configuration of external inputs.

2.2.5 Boolean modelling formats and tools

A BM can be constructed and represented using various modelling tools relying on different formats, as illustrated in the (Figure 2.1). One of these formats is the SIF [127], which is used for encoding a model topology from a list of interactions, giving an easy solution for combining new interactions to models. SIF is supported by different tools and databases such as Cytoscape [128], OmniPath [129] and Signor [130].

In order to re-use or integrate models, they need to be translated from their original format. Literature-constructed diagrams can be transformed into BMs as a SIF using automated conversion tools (Figure 2.1). SIF can be translated into a list of BFs in Boolsim format [131] using a Standardised QUAlitative Dynamic approach (SQUAD) [132]. In addition, GNA allows the encoding of model functions and specifies qualitative values of a model from the experimental literature [133].

Model annotations can be stored along with the topology and BFs using a SBML-qual. SBML-qual is a standard format designed by the CoLoMoTo com-

munity [134], extending the SBML [135] to represent the qualitative models of biological networks [136]. Pathway diagrams from KEGG, BioCarta and SABIO-RK [137] can be transformed to SBML-qual using PAth2Models [137]. Using a dedicated converter CaSQ [138], CellDesigner SBML format can be translated to SBML-qual. Notably, SBML can be translated into SBML-qual by CellNOpt[139] and SQUAD [132]. However, the SBML-qual format is still incompatible with some tools such as RMut [122], NetDS [140] and CABEAN [141], pointing out that further integration efforts are required to allow reproducibility of SBML-qual models with the incompatible tools.

In the modelling process, a range of tools is available for model inference, simulation, and attractor analysis. In the case of model inference, the majority of represented tools in the (Table 2.1) generate BMs randomly based on selected information, which is known as random model generation. The two software-CABERNET and caspo- create a BM by augmentation if the topological/ functional characterization is incomplete (Table 2.1). A user can generate ensembles of models by combining components and interactions with the original model. The synchronous updating scheme is a default simulation method, supported by all the reviewed tools. Moreover, CellCollective [142], GINsim [143], ADAM [144], BoolNet [102], MaBoSS [145] and CABEAN [141] support the asynchronous scheme (Table 2.1). Most of the represented tools identify the attractor dynamics with heuristic and exhaustive search except BooleSim [89], CellNopt, RMut and SQUAD. Once the attractors are identified, the stability and controllability check can be performed by RMut, CANA [146], CABEAN, ASSA-PBN [147], BioModel Analyzer BMA [148]. In turn, BoolNet, RMut, GDSC [149] and CABERNET[150] perform the topological analysis of the intrinsic structure of the model. As mentioned, particular tools such as RMut, NetDS, and CABEAN are incompatible with the SBML-qual modelling format. There may be limitations to the reusability of models created with these tools since they have their own formats. As a result, ensuring interoperability and reproducibility of models is necessary when incompatibilities exist.

2.3 Applications of Boolean modelling in clinical and translational medicine

Boolean modelling was applied in clinical and translational medicine research[34, 35, 151, 36] for various purposes. Simulation of the complex biological systems allowed to predict the activity of pathway endpoints (phenotypes) [152], drug targets[125] and crosstalks[153]. Identifying attractors helped to understand the activity of the phenotypes, since they represent the steady states of

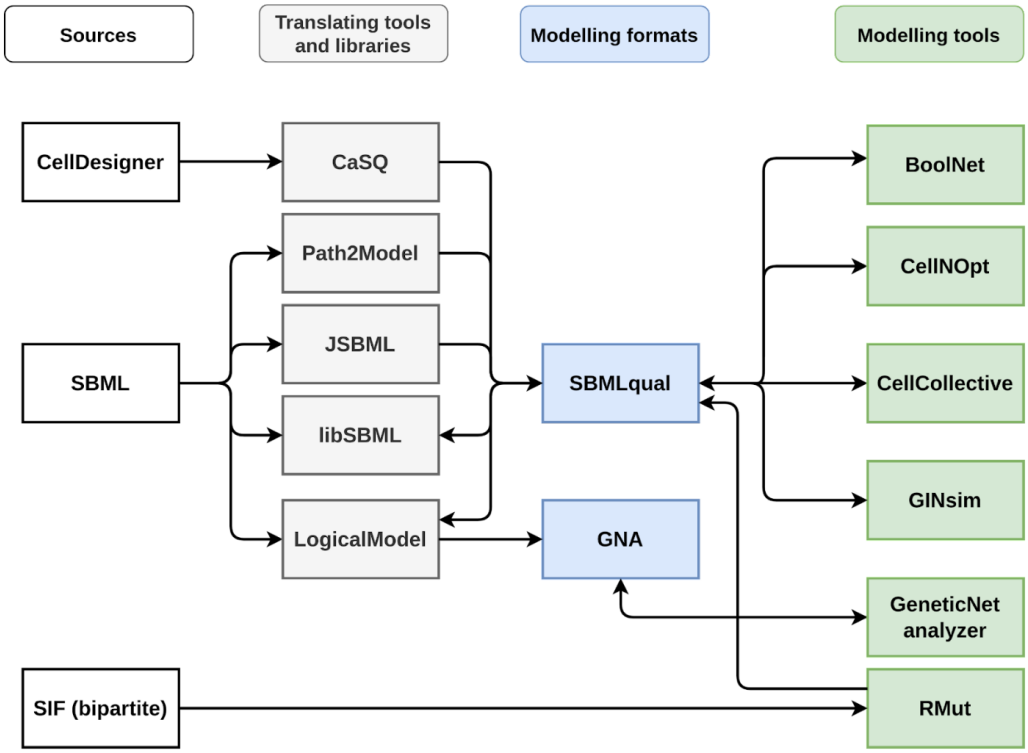


Figure 2.1 Overview of the interoperability of Boolean modelling tools, libraries and formats. The format of data resources (white colour) can be translated by tools and libraries (grey colour) to modelling formats (blue colour), to be used by the popular Boolean modelling tools (green colour).

Tool	Interface	Format	Model generation	Updating scheme	Attractor search	Attractor analysis	Topological analysis
CellCollective	Web, GUI	SBML qual	–	Asynchronous	Heuristic, Exhaustive	–	Centrality
GINSIM	GUI	SBML qual	–	Asynchronous	Heuristic, Exhaustive	–	–
Boolesim	Web, GUI	Own	–	Synchronous	–	–	–
ADAM	Web, GUI	SBML core	–	Asynchronous	Heuristic, Exhaustive	–	–
BoolNet	CL, Cytoscape	SBML qual	Random	Asynchronous	Heuristic, Exhaustive	–	Centrality, Clustering
CellNopt	CL, Cytoscape	SBML core, SBML qual	–	Synchronous	–	–	Centrality
RMut	CL	Own	Random	Synchronous	–	Stability, Controlability	Centrality, Clustering
SQUAD	GUI	SBML core, SBML qual	–	Synchronous	–	–	–
CABERNET	GUI, Cytoscape	SBML core	Random, Augumented	Synchronous	Heuristic, Exhaustive	Stability	Centrality, Clustering
NetDS	GUI	SBML core	Random	Synchronous	Heuristic, Exhaustive	Stability	Centrality
GDSC	Web, GUI	Own	–	Synchronous	Heuristic, Exhaustive	–	Centrality
CANA	CL	Own	Random	Synchronous	Exhaustive	Stability, Controlability	–
CABEAN	CL	Own	–	Asynchronous	Exhaustive	Stability, Controlability	–
ASSA-PBN	CL	Own	Random	Synchronous	Heuristic, Exhaustive	Stability, Controlability	–
caspo	CL	Own	Augumented	Synchronous	–	–	–
BMA	Web, GUI, CL	Own	–	Synchronous	Exhaustive	Stability, Controlability	–

Table 2.1 Summary of key tools and their functionalities that were implemented to perform Boolean analysis and simulations. GUI - Graphical User Interface, CL - Command Line.

biomolecules[100, 154]. Finally, comparing attractors before and after perturbations allowed evaluating the model stability and gave insight into how the in-vivo systems maintain their homeostasis. Below examples of such applications are discussed, and summarised in (Table 2.2).

2.3.1 Modelling of cell signalling

A complex signalling network can determine cellular decisions, but the kinetic parameters and quantitative data that enable dynamic modelling may not be sufficient. Therefore, computational approaches based on the qualitative structure of these networks are of great interest. Boolean modelling of cellular signalling provided insights into the process of signals and the interactions between regulators and target molecules. A model of T cell signalling included: i) a T cell receptor, its co-receptors CD4/CD8, and CD28 which regulates T cell function, ii) selected MAPK signalling and PI3K/PKB signalling driving cellular activation and differentiation. The model was able to reproduce the literature and experimental results upon different activation scenarios of TCR, CD4 and CD28. Moreover, it reproduced the T cell phenotype in response to knockouts and predicted unexpected activation of the PI3K/PKB pathway after TCR activation [155]. This model was extended [156]into large BMs modelling a regulatory Th cell. TCR signalling, cytokine signalling, and cell cycle models were studied separately, and integrated into a single model. The model showed the naive cell differentiation into Th1, Th2, Th17 and Treg subtypes. The analysis predicted an unexpected plasticity behaviour of the canonical cell types as well as the potential of regulatory T cells to differentiate into Th1 or Th2 subtypes.

Another BM modelled the plasticity of CD4+ T cell differentiation [157] and showed that it is controlled by the dose and composition of cytokines. The model explained the T cell fate by defining 500 external conditions and considering all possible endogenous interactions. These interventions were perturbed to control the dynamics of the model from undesired to desired phenotypes. The model reproduced known synergistic actions of feedback loops on IL-12R expression and confirmed results from other studies [158, 159], showing that the balance between i-Treg and Th17 was regulated by IL-6. Furthermore, the model predicted a complex phenotype (Th1-Th2) after activation of Tbet and GATA3 transcription factors under the similar environmental conditions proposed by an in vivo study [160].

Integrating different layers of biological data allowed for understanding the heterogeneity of multifactorial diseases and for reducing the possibility of false

positive results. A BM was used to analyse the regulation of key transcription factors (TFs) in Rheumatoid Arthritis (RA) and derive patient-specific models to understand the disease complexity and the response to treatment [161]. The model highlighted the impact of TGFB1, IL6, and TNF in response to the anti-TNF drugs on the model outputs. The analysis showed that TFs are master regulators- the activation of IL6 and/or TGFB1 positively regulates TFs expression, even with deactivation of TNF cascade. Blocking IL6 and TGFB1, and TNF cascades deactivates TFs expression. Further, the MAPK molecules depend on the activation of IL6 and TGFB1 and do not be affected by TNF deactivation.

2.3.2 Modelling of cancer growth signalling and apoptosis

In cancer, targeted therapies inhibit driver molecules of tumorigenic pathways [162]. However, it is difficult to identify targets that have crucial functions in tumour progression because of complex interactions and feedback loops between implicated molecules. Moreover, monotherapies were found to be additive in their actions because tumours are highly complex and evolve continuously [163]. Therefore, they had limited efficacy and needed many clinical trials. Perturbation analysis may help to understand this complexity, proposing the interventions between molecular targets and predicting their possible synergistic action. The BMs of gastric cancer used this analysis on seven known inhibitors that target the gastric cancer pathways [35]. All possible combinations were calculated then simulated in silico to identify new synergistic targets, which were then experimentally validated. In another work, probabilistic BM allowed associating the activity of the pathological phenotype to the perturbation probability of its regulators. Under a given perturbation, the model tested the possible synergistic perturbations to decrease the activity of the phenotype [164].

Boolean modelling was proposed to simplify the complex interactions and their downstream signals. The molecular intervention analysis showed that the combinatory inhibition of oncogenic molecules e.g. PDK1, AKT, and MDM2 or the activation of P53, RB and CDH1 reduces the proliferation and increases quiescent phenotypes since the targeted drug associations blocked cancer pathways at different regions [162]. Signalling networks in cancer are complex cascades and their pathological rewiring may alter cellular proliferation, migration and apoptosis resistance [165], and BMs can help to understand this complicated rewiring [166]. A BM was constructed combining the main cancer pathways such as RTKs, Wnt/β-catenin, TGF-β/Smads, Rb, HIF-1, p53, PI3K/AKT signalling pathways [165]. Identified attractors were associated with apoptosis, proliferation, and quiescent phenotypes in response to environmental conditions. The model revealed that growth factor signalling significantly increased the proliferation

and quiescent phenotypes but decreased the apoptosis. The similar result was proposed by another model [167]which combined the intrinsic and extrinsic pro-apoptotic pathways with the growth factor signalling.

In another study, a BM describing the PI3K/AKT1 signalling pathway showed increased tissue proliferation and cell invasion phenotypes [115]. In particular, the oscillations of PI3K protein expression were studied by simulating its different activity levels at different cellular stages. Using different updating schemes can be more appropriate in specific settings and this is an example that illustrates it - While applying the synchronous updating scheme, the inhibition effect of PI3K induced four phenotypes including G2 arrest, mitotic catastrophe, and aberrant and normal anaphase. However, the asynchronous scheme showed that the previous four phenotypes didn't occur at the same time, and they are not synergistic in signal transduction because the asynchronous scheme updates the biomolecules at different time intervals. Therefore, depending on the biological process and the knowledge about the real biological time, modellers get to decide which updating schemes make more sense to achieve a desired output.

Logical models of cancer are usually generic because they use heterogeneous data and require clinical data to calibrate them. To generate precise BMs, a PROFILE framework [165] was proposed, integrating the mechanistic insights of logical modelling with multi-omics data. The PROFILE framework combined mutations and expression data (METABRIC [168], TCGAdataset https://www.cancer.gov/tcga) with the cancer BM to simulate different cases and compare the model outputs. After data binarization, the activity of the nodes and the transition rates were modified based on specific cases. Stochastic simulations were performed using MaBoSS [169] for a semi-quantitative analysis of model perturbations. This approach was used in another study to investigate BRAF inhibition in melanoma and colorectal cancer which have significant variations despite the similar omics profiles [37]. The model was able to differentiate between the two cancers based on different datasets. This approach extends the previous works using the dynamic data [170] and the same pathways [171]to personalise the signalling behaviours in response to treatments.

Recently, researchers tested the PROFILE framework on a prostate model and infer patient-specific treatments [172]. The model of prostate cancer includes major deregulated signalling pathways integrated with mutation and RNAseq data. The biomolecules are fixed to zero/one according to the type of the mutations. For the continuous RNAseq data, the expression levels are translated as a modulation of a signal to the initial conditions to influence the probability of transitions. The

Model	Nodes	Interactions	Type	Ref.
T cell signalling (MAPK signalling and PI3K/PKB signalling)	94	123	Cell signalling	[173]
TCR signalling, Cytokine signalling, and cell cycle	65	135	Cell signalling	[156]
Plasticity of CD4+ T cell differentiation	38	96	Cell signalling	[174]
TGFB1, IL6, and TNF signalling	38	59	Cell signalling	[161]
Gastric adenocarcinoma	10	34	Cancer signalling	[175]
Simplified cancer network	96	249	Cancer signalling	[176]
RTKs, Wnt/β-catenin, TGF-β/Smads, Rb, HIF-1, p53, PI3K/AKT signalling pathways	98	254	Cancer signalling	[165]
Pro-apoptotic pathways with the growth factor signalling	37	63	Cancer signalling	[177]
PI3K/AKT1 signalling pathway	30	42	Cancer signalling	[178]
Signalling pathways around BRAF in colorectal and melanoma cancers.	33	43	Cancer signalling	[37]
Signalling in prostate cancer	133	449	Cancer signalling	[172]

Table 2.2 Selected models and applications of Boolean modelling in clinical and translational medicine, with an overview of their scale in nodes (graph vertices) and interactions (graph edges).

analysis highlights that apoptosis is activated by Caspase 8/9, while proliferation is activated by cyclins D/B. Further, several readouts of cancer hallmarks (phenotypic outputs) were detected such as metastasis and DNA repair. The analysis identifies a list of drug combinations that reduce the proliferation phenotype or increase the apoptosis. The researchers use Boolean simulations to grade the effect of the combined drugs on patient-specific phenotypes, comparing the effects of treatments on each patient to suggest suitable treatment.

Chapter 3

Methods

This Chapter describes methods for construction of BMs from systems biology diagrams, and their subsequent simulation. In particular, the conversion from CellDesigner SBML graphical format is discussed, and the approaches to validate the resulting models. Next, omics data used for creation of cohort-specific models are discussed. Finally, the framework for stochastic simulation of such calibrated models is described. Discussed methods and data are available in the LCSB GitLab repository.[1]

3.1 Construction of Boolean models

3.1.1 Construction of BMs from systems biology diagrams

Systems biology diagrams in the Process Description format (see Section 2.2.1) were used to construct Boolean models. Diagrams in CellDesigner SBML format were obtained from the Parkinsons's disease map hosted on the MINERVA Platform [15]. The MINERVA Platform allows to export selected parts of the map. Such parts are henceforth called diagrams.

3.1.2 Translation of the diagrams into Boolean models using CaSQ

The diagrams in CellDesigner SBML format were translated to SBML-qual using CaSQ (CellDesigner as SBML-qual) automatically. The translation process starts with the diagram reduction based on specific rewriting rules [93]. These steps are illustrated in Figure 3.1.

[1] `https://gitlab.lcsb.uni.lu/Ahmed7emdan/Boolean modelling of PD -Thesis`

First step of this translation process is the reduction of the Process Description notation into Activity Flow notation (see Section 2.2.1). Biomolecules in Process Description diagrams can be represented in several different states (e.g. phosphorylated, methylated etc) and upon translation into Activity Flow they are represented a by a single node in the diagram, with the different states of the biomolecule represented by different logical states of the node (e.g., active or inactive).

CaSQ also infers the logical functions and translates the interactions from the Process Description to the Activity Flow, considering one qualitative species to each species after the reduction step. A Process Description is a detailed description of the steps or events that occur during a particular process or mechanism. It provides a detailed account of how something happens or works. An Activity Flow, on the other hand, is a visual representation of the steps or events in a process, showing the sequence of activities and the relationships between them in a clear and concise manner. In the context of biological mechanisms, a Process Description describes the various molecular reactions and interactions that occur during a particular process, such as DNA replication or protein synthesis. An Activity Flow depicts the same process as a series of interconnected steps, showing how one activity leads to another and how different components of the system interact with each other. One way to turn a Process Description into an Activity Flow interaction is to represent each step or event in the process as a node in a diagram, and to represent the relationships between the steps or events as connections or edges between the nodes (see Figure 3.1). In order to facilitate the use of BMs with different tools, the diagrams were transformed to the Simple Interaction Format SIF using CaSQ tool.This process allowed for the creation of SBMl-qual models that could be utilised with certain tools such as BoolNet (Table 2.1).

Rule 1 In a reaction, where a receptor and a ligand form complex, the receptor reactant is removed.

Rule 2 In a reaction, where two proteins form a complex, the reactants are removed and modifiers are linked directly to the product.

Rule 3 The inactive forms of a biomolecule in a single reaction is removed (do not participate in other interactions)

Rule 4 In a single transport reaction a reactant is removed and the reactions are linked to the product (in case product and reactant are the same).

3.1.3 Evaluation of biological relevance

The accuracy of representations of a biomolecular mechanisms were evaluated. This included:

First: Comparison with exploratory literature: Conduct an exploratory literature search to determine the validated effects of molecular perturbations in vivo. To do so, it is important to first define specific key biomolecules using the structural and sensitivity analysis. This helped to guide the selection of appropriate dataset and ensure that the studies reviewed are relevant. Once the key molecules are defined, the next step is to identify and select appropriate databases to search. There are many databases available that contain information on scientific and medical research, such as PubMed and Scopus. It is important to choose databases that have a reputation for publishing high-quality research. In addition to choosing reputable databases, it was also important to use filtering criteria to ensure that the reviewed studies are of high quality. Some common filtering criteria that were considered include the study design (e.g. randomized controlled trials), the sample size, the duration of the study, the type of intervention being tested, and the type of outcome being measured. For some articles that show discrepancies, some tools such as the Cochrane Risk of Bias tool [179]or the Newcastle-Ottawa Scale [180] were considered to assess the quality of the reviewed studies. The impact factor (IF) of a journal was also considered. However, it is important to keep in mind that the impact factor is not a perfect measure of the quality of individual articles or the importance of the research they contain. The impact factor is based on citations, which may not always be an accurate measure of the quality or importance of a piece of research. Therefore, it was important to use the impact factor in conjunction with other criteria.

Once the appropriate databases have been selected, the next step is to use keywords related to the research question to search for relevant studies. This can be done by using advanced search features within the database or by using logical operators to narrow the search results. Once a list of relevant studies was identified, it is important to review the abstracts of these studies to determine their relevance. If a study seems relevant, one can get a more complete understanding of the research and its findings. It is also important to create a dedicated online library, recording the title, author, publication year, and a brief summary of the study, as this information will be useful when analyzing and interpreting the results.

Once all relevant studies are gathered, it is important to analyze the results and interpret the findings. This may involve looking for patterns or trends in the literature and considering how the findings relate to the original research question. It is also important to consider any gaps in the current knowledge about

the topic and to identify areas where further research is needed.

Second: Comparison with omics data: the first step was to identify the model biomolecules and their molecular functions. Then, a list of differentially expressed genes was obtained, either by performing differential expression analysis on a dataset or by finding a published list. The differentially expressed genes are then mapped to the model biomolecules using enrichment analysis in PD map and other resources (Section 2.2.1). Finally, the overlap between the two lists was analyzed and statistically tested to determine if it is significant.

3.2 Topological analysis of the models

The interactions between biomolecules in a BM was analysed to study their structure and correctness. The topological features of the BM as a network were analysed to understand the relationships and interactions between biomolecules. These interactions were represented as a graph, with the biomolecules as nodes and the interactions between them as edges. Statistical packages were used [127] to calculate the directionality and the degrees of the nodes (in degree/out degree), feedback loops, and centrality measures. Centrality measures, such as betweenness and stress were used to quantify the importance or influence of a particular node in the network[181]. Betweenness centrality measures the number of times a node lies on the shortest path between two other nodes, with nodes that act as intermediaries between other nodes considered more central. Stress centrality is a measure of a node's importance in a network based on the number of shortest paths that pass through it. A node with a high stress centrality is considered more central because it plays a key role in connecting other nodes in the network, as it is part of many shortest paths. This type of centrality is useful for identifying nodes that are critical for the overall connectivity of the network [182]. These measures were used to identify key nodes in the model, such as hubs and to understand the role that these nodes play in the overall model. They were used to identify potential targets for therapeutic interventions, as targeting highly central nodes may have a greater impact on the overall model.

In order to use the topological analysis to validate the correctness of the models, certain criteria were considered based on the biological knowledge of the systems being modelled and the centrality measures of the key biomolecules. Based on the topological analysis, the top five centrality measures were selected to search for biological relevance Tables 4.2 and 4.6.

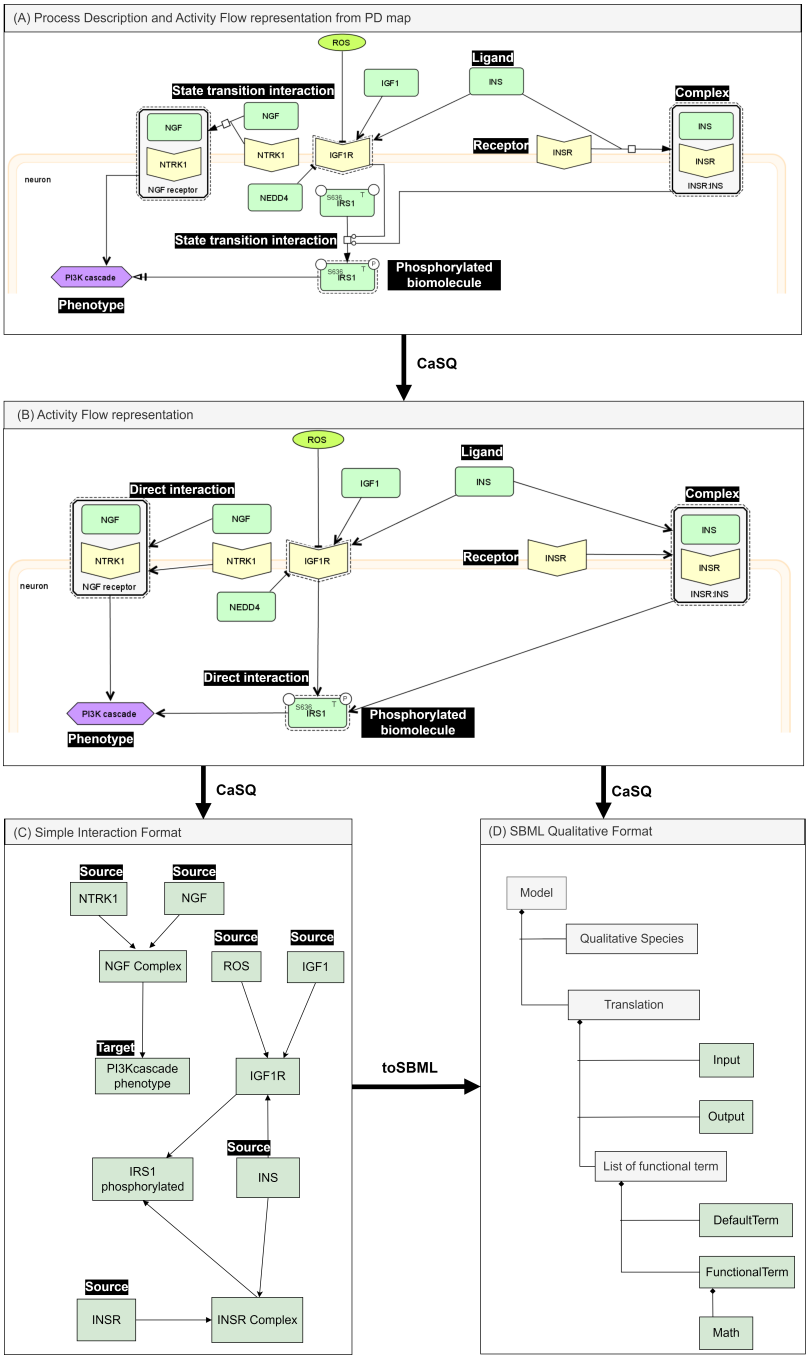


Figure 3.1 Translation from a mixed representation of Process Description and Activity Flow, to an Activity Flow representation, and finally to a Boolean model. Panel A shows the Process Description representation, featuring indirect state transition interaction. Panel B illustrates the transformation from state transition to direct interaction in the Activity Flow representation. Panel C displays the Simple Interaction Format (SIF) representation, with sources and targets. Panel D indicates the basic structure of SBML-qual, generated from either the Activity Flow or the SIF

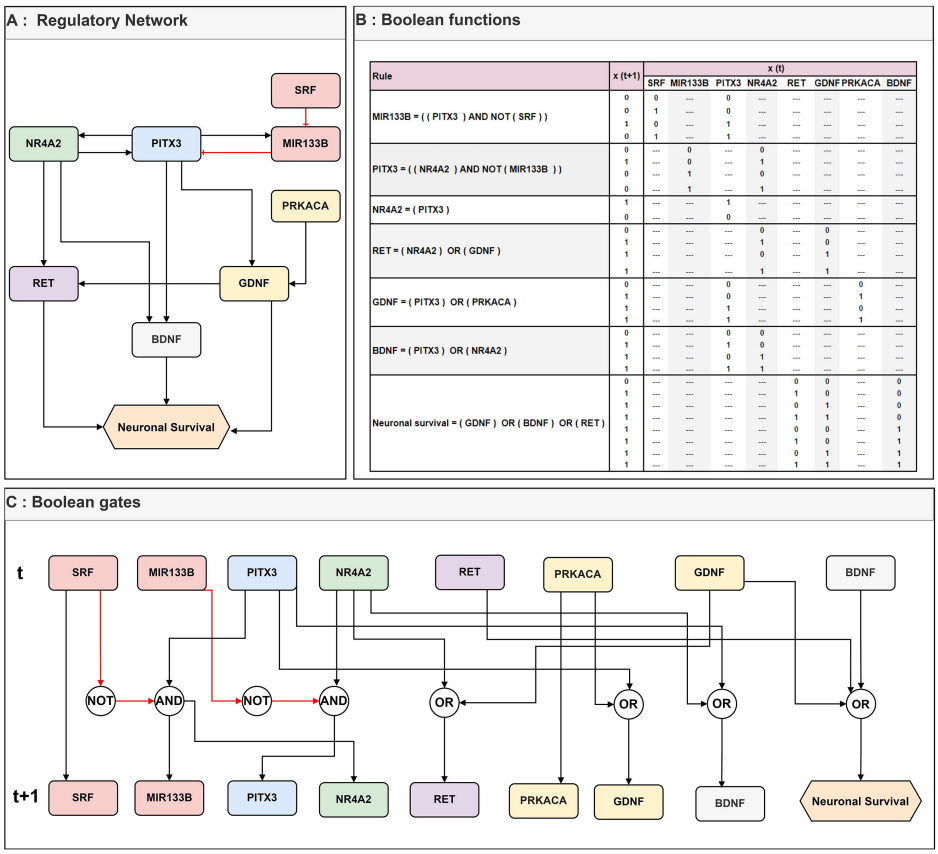


Figure 3.2 (A) illustrates a simple directed Network [35], with typically used logical functions. Red arrow refers to the inhibition effects. Black arrows refer to the activation effect. (B) shows Boolean functions either in basic logical expressions or as a truth table. (C) shows the Boolean gates AND/OR/NOT, describing the dynamics update from time (t) to (t + 1).

3.3 Model analysis

3.3.1 Model updating schemes

Different updating schemes were compared and evaluated. The BFs were updated by synchronous and asynchronous updating schemes and a hybrid scheme that combines both approaches [98] (see Figure 3.4).

The synchronous updating scheme changes the state of all biomolecules at the same time, according to their BFs. It is deterministic, meaning that the same initial conditions will always result in the same final state. This makes it easier to predict the behavior of the system and to understand how changes in the BFs of the biomolecules will affect the system's behavior. However, synchronous updating can be difficult to model systems with time delays or systems that have multiple, interacting sub-systems, as the synchronous update scheme does not allow for different biomolecules to update at different times.

In contrast, the asynchronous updating changes states of biomolecules at different times based on their BFs. This can make it easier to model systems with time delays or complex interactions between biomolecules. However, it can also be more difficult to predict the behavior of the system, as the same initial conditions can result in different final states depending on the order in which the biomolecules are updated[100].

These schemes can be updated in probabilistic manner using probabilistic BMs. Probabilistic BMs assigns certain probabilities to BFs, and each biomolecule is updated based on this probability before reaching the steady state [101, 104].

These schemes are summarised in Figure A (Figure 3.4), BM includes three components X1, X2, X3 which have states (zero/one). The dynamics of a component is represented by Boolean function. Synchronous updating scheme updates all states at the same time, the successor states have two possible values, one (ON) or zero (OFF). In the asynchronous updating scheme, the start states are not updated at the same time (one state is updated per iteration), the successor states have two possible values one (ON) or zero (OFF). In Figure B (Figure 3.4), a probabilistic BMs shows that states are updated at the same time and the successor states present different probabilities; p represents the updated probability values of the variables. Importantly, an asynchronous updating scheme can be used in probabilistic BMs as well.

To visualize the behavior of a BM under different update schemes, state transition graphs were generated. This is a graphical representation of all the possible states that the system can be in, along with the transitions between states. The state transition graph can be used to illustrate the spectrum of all possible outcomes for a given initial condition, depending on the update scheme being used. For further examples of state transition graphs, see the "steady states

calculations" directory LCSB GitLab repository. [2].

In this study, the performance of synchronous and asynchronous update schemes were evaluated using various metrics to assess the stability and predictability of the models. These metrics included tracking the number and duration of stable states or fixed points produced by each update scheme. The results of this comparison are presented in the "steady states calculations" directory in the GitLab repository [183]. Both updating schemes were able to replicate known or expected behavior in the system. However, the choice of which update scheme is superior depends on the specific biological process being modeled and the available knowledge about the real biological time course. Asynchronous update schemes may take longer to achieve a desired output, but may be more biologically realistic. On the other hand, synchronous update schemes may be faster, but are deterministic in nature.

3.3.2 Attractor Search

Attractors are states in a state transition graph with no outgoing edges. These states can be classified into two categories: stable states (also known as fixed points), which do not change over time, and complex attractors, which are sets of possible outcomes that can be reached through synchronous or asynchronous updating schemes[184]. The set of states that lead to an attractor is known as the basin of attraction, which can be interpreted as a set of potential biological scenarios that can be tested through hypotheses [185].

In this study, several search algorithms were compared during model verification by evaluating their speed and reachability (Table 4.3). An exhaustive search was employed, in which all possible attractors were found through synchronous transitions between states. To improve the speed of this approach, a SAT-based method was used to formulate the attractor identification problem as a Boolean satisfiability problem (SAT). This allowed for the determination of whether a given formula was satisfiable or unsatisfiable, and for the restriction of the search to loops of a specified length [110].

In comparison, the decomposition method is a computational approach that aims to optimize speed and reduce complexity by dividing a model into strongly connected components (SCCs). While this method may be effective at improving the speed of getting attractors, it may not always accurately reflect biological processes. This is because the primary focus of the decomposition method is on optimizing for speed, rather than accurately modelling the underlying biological processes. As a result, the decomposition method may oversimplify the model and lead to inaccurate results in certain contexts.

[2] `https://gitlab.lcsb.uni.lu/Ahmed7emdan/Boolean modelling of PD -Thesis`

Figure 3.3 An example of a state transition graph for a dopamine transcription Boolean model under a synchronous update scheme. The illustrated nodes are attractors, showing the different possible steady states in red and the transitions that lead to them in green. This graph illustrates the range of outcomes that can occur based on the initial conditions and update scheme used.

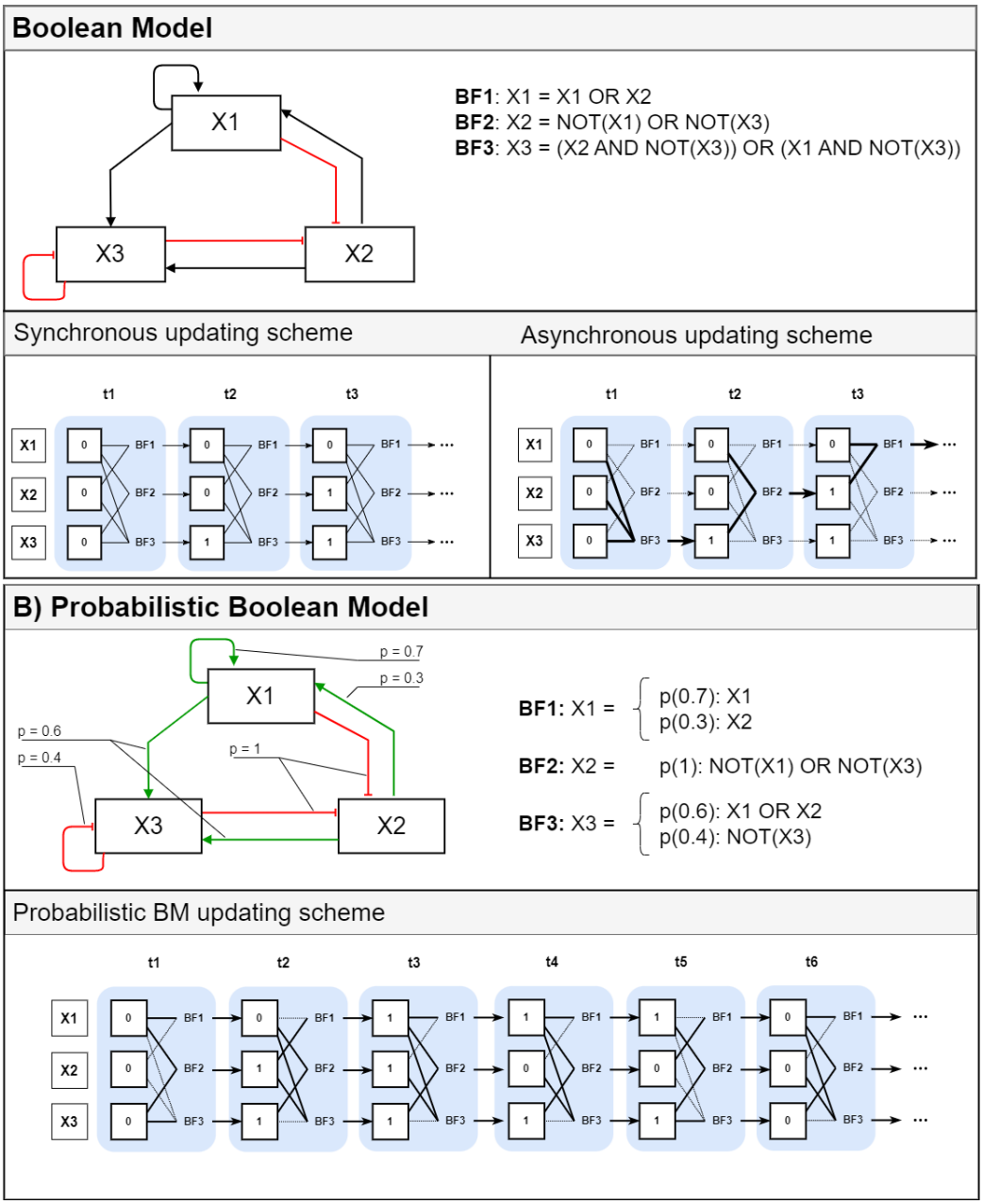


Figure 3.4 The figure represents the model updating in time for a simple regulatory graph. (A) Boolean model with three components X1, X2, X3 which have states (zero/one) and BFs. (B) A Probabilistic Boolean model shows that states are updated at the same time and the successor states present different probabilities.

In addition to the exhaustive search, heuristic and asynchronous searches were also conducted. For the heuristic search, a subset of possible states was specified as initial conditions and synchronous transitions were performed until an attractor was reached. For the asynchronous search, random asynchronous transitions were used to identify steady states and complex attractors from the specified initial conditions. This resulted in the network's current state being located within an attractor with high probability, and allowed for the identification of forward reachable states and the comparison of these to the current state sets in order to identify complex attractors [186].

The results and visualisation of the attractor analysis, including details on steady states and complex attractors, can be found in the "steady states calculations" directory within the LCSB GitLab repository [183].

3.3.3 Perturbation analysis

Perturbation analysis was used in order to assess the impact on the topological robustness and dynamic resilience of the models, as well as the attractors reached by them. [127] (Section 2.2.4). Two main types of perturbations were evaluated (see also Figure 3.5):

- node perturbations, which involve altering the state of a single biomolecule. These were performed through knockout or overexpression.
- edge perturbations, which involve altering the function of the interactions between biomolecules. These were performed by edge removal or attenuation mutation.

The evaluation was based on performing sensitivity analysis on each biomolecule and interaction in a set of selected models. The sensitivity analysis is a technique used to assess how the output of a model or system changes in response to perturbations, i.e. two attractors reached by a given model (unperturbed and perturbed). The distance between two attractors, was measured using similarity-based distance and identity-based distance . Similarity-based distance measured the similarity between two attractors, taking into account the common and unique states present in both attractors. Identity-based distance, on the other hand, measured the percentage of states that are present in both attractors

After the evaluation, only node-based perturbations (knockout and overexpression) were considered. Changing the original rules of the model through perturbations such as rule-flip or edge perturbations significantly affected the stability of the models, resulting in large changes in their attractors. This made it difficult to accurately assess the impact of the perturbations on the model, as the

large changes in the attractors were not representative of the true behavior of the model. The results of the selected perturbation methods- node-based knockout and overexpression can be found in Figures 4.1 and 4.2.

The results of sensitivity analysis of other types of perturbations including the edge perturbations are available in the GitLab repository [183]. The data include the equations and codes used to calculate the sensitivities.

3.4 Integration of Boolean models with omics data

The section involves the use of two datasets: the Parkinson's Progression Markers Initiative-microRNA (PPMI-miRNA) dataset and the Type 2 Diabetes Mellitus (T2DM) dataset. The PPMI-miRNA dataset includes miRNA expression data collected as part of a clinical study on Parkinson's disease, and in this section, the differential expression of miRNAs is analyzed and the identification of miRNA targets is performed. The T2DM dataset consists of transcriptomic data on PINK1 and GBA mutations in T2DM, and this data is analyzed to examine the specific impact of perturbation of T2DM on PD progression. Finally, the section explain the use of the pyMaBoSS framework to perform stochastic Boolean model simulation.

3.4.1 Parkinson's Progression Markers Initiative-miRNAs dataset

The PPMI (Parkinson's Progression Markers Initiative) dataset is a multi-cohort, longitudinal observational study that examines the molecular and clinical changes in different subtypes of Parkinson's disease (PD). The dataset includes microRNAs from blood samples collected from individuals with clinical PD, prodromal PD, scan without dopaminergic deficit (SWEDD), and Parkinsonism. Clinical PD refers to individuals who exhibit the clinical symptoms of PD and have a positive dopamine transporter (DAT) SPECT scan. Prodromal PD refers to individuals who do not yet have severe symptoms but do have significant positive DAT SPECT results (223 individuals). SWEDD refers to individuals who have been clinically diagnosed with PD but do not show dopaminergic deficit on their DAT SPECT scan (187 individuals). Atypical Parkinson's disease, or parkinsonism, refers to individuals with idiopathic symptoms similar to those seen in typical PD. This data is available through the LONI (Laboratory of Neuro Imaging) archive at www.ppmi-info.org/data.

Diagrams used in the stratification process are selected based on pathway enrichment analysis of the Parkinson's Progression Markers Initiative (PPMI)

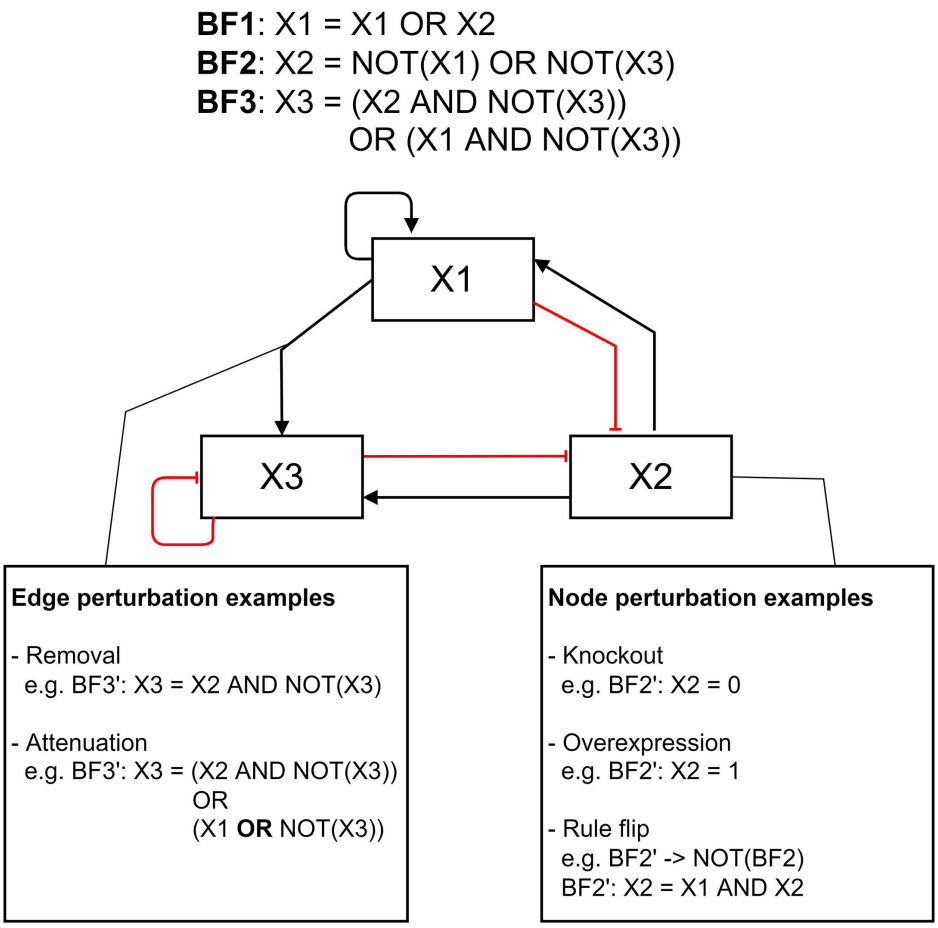


Figure 3.5 The figure illustrates a regulatory graph in which the states of nodes X1, X2, and X3 can be modulated by activating (black link) or inhibiting (red link) influences. Node perturbations represent changes in the states of these nodes that occur due to knockout or overexpression. Edge perturbations depict alterations in the functions of the interactions between the nodes resulting from mutations.

dataset using the Parkinson's disease map (PD map). Once significant pathways are identified, they are exported in bipartite and CellDesigner SBML formats for further modelling. Both formats are translated to SBML-qual, which is a module of the SBML standard that is specifically designed for representing qualitative models of biological systems.

Statistical correlations and Effect size of PPMI dataset

The differential expressed miRNAs were calculated by estimating the Log2 fold change between cohorts(case/control). Additionally statistical correlations and effect sizes were calculated.

Statistical correlations: Moreover, statistical correlations are calculated including paired sample t test, Pearson correlation and Wilcoxon rank test.

Paired t-test: The paired t-test is a statistical test used to determine whether there is a significant difference between the means of two related groups [187]. It is based on the t-statistic, which was calculated as follows:

$$t = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}.$$

where:

- $\bar{x}1$ is the mean of the first group
- $\bar{x}2$ is the mean of the second group
- s1 is the standard deviation of the first group
- s2 is the standard deviation of the second group
- n1 is the size of the first group
- n2 is the size of the second group

Wilcoxon rank sum test: The Wilcoxon rank sum test is a non-parametric statistical test used to determine whether there is a significant difference between the medians of two groups [188]. It is based on the rank sum statistic, which was calculated as follows:

$$W = \sum R_i$$

where:

- Ri is the rank of the ith observation in the combined dataset

Pearson's correlation coefficient: It is a measure of the strength and direction of the linear relationship between two variables [189]. It was calculated as follows:

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2}\sqrt{\sum (y_i - \bar{y})^2}}$$

where:

- xi is the ith value of the first variable
- yi is the ith value of the second variable
- $\bar{x}$ is the mean of the first variable
- $\bar{y}$ is the mean of the second variable

Pearson's correlation coefficient can range from -1 to 1. A value of 1 indicates a perfect positive correlation, meaning that as the value of one variable increases, the value of the other variable also increases. A value of -1 indicates a perfect negative correlation, meaning that as the value of one variable increases, the value of the other variable decreases. A value of 0 indicates no correlation.

Effect sizes: The effect sizes of each miRNA were identified by calculating the Cohen distance between multiple cases[190]. Cohen's distance is a standardized measure of the difference between two groups' means divided by the data's standard deviation. This measure calculates the size of the mean difference by comparing it to the pooled standard deviation. One of the advantages of the Cohen distance is that it can be translated into probabilities using Common language effect size (CL) [191, 192]. The probability describes how a random outcome observation from one group differs from a random outcome observation from the other group. Using this method, Cohen distance information is communicated in a more intuitive manner[191]. The Cohen distance is estimated as follows:

$$d = \frac{\bar{Y}_1 - \bar{Y}_2}{s_{\mathrm{p}}},$$

where $\bar{Y}_1$ and $\bar{Y}_2$ are the sample means and s_{p} is the pooled standard deviation; $= \sqrt{p_1 s_1^2 + p_2 s_2^2}$, where p_1 and p_2 are the base rates in the sample ($p_1 + p_2 = 1$) and s_1^2 and s_2^2 are the sample variances.[191, 192]. CL is calculated using the following formula:

$$\mathrm{CL} = \Phi \cdot \left(\frac{\bar{Y}_1 - \bar{Y}_2}{\sqrt{s_1^2 + s_2^2}} \right),$$

where Φ is the cumulative normal distribution function [192].

Identification of the miRNA targets

The validated targets of miRNAs are identified using manually curated databases such as miRTarBase[193]. This database contains more than three hundred and sixty thousand experimentally validated miRNA-target interactions, which are systematically identified through natural language processing of the text to filter articles relating to functional studies of miRNA.

The predicted and validated miRNAs-targets are enriched using PD map as a disease-specific knowledge resource to identify the overlaps with substatia nigra targets, making miRNAs more specific to PD and identify the involved pathways. The substantia nigra dataset consists of microarray gene expression data for human post mortum brain samples from various regions, human whole blood samples, samples from animal models, and cell culture samples. This data is related to PD map and is publicly available [194].

The expression pattern of the miRNAs are compared with the curated miRNA expression databases: MiREDiBase[195], miRGate[196], The Human miRNA and Disease Database (HMDD)[197], and GEO data screening[198] .

3.4.2 Identification of Type two Diabetes mellitus transcriptomic profile

Type two Diabetes mellitus (T2DM) is a chronic condition characterized by high levels of sugar (glucose) in the blood which can be a comorbidity of the PD. To investigate the relationship between the T2DM on the progression of PD subtypes identified in PPMI the in-house unpublished transcriptomic datasets were used that describe PINK1 and GBA mutations in T2DM. These datasets were generated as a part of an internal LCSB project to study T2DM-PD comorbidities. The publication is being prepared and the datasets are available upon request.

In the first dataset the samples were analysed from three distinct PD patients carrying the homozygous Q456X mutation in PINK1 (compared with their corresponding isogenic gene-corrected controls). The dataset comprise a list of RNA-seq differentially expressed genes analyzed from iPSC-derived neurons (30 days of differentiation). In this dataset, the relationship between PD and T2DM is established based on the following findings. It was observed in vivo that IRS1 levels were decreased and phosphorylation of AKT at both S473 and T308 was reduced in PINK1 mutant neurons compared to isogenic controls, indicating impaired insulin signaling.

In the second dataset, differentially expressed genes of the GBA N307S mutation were identified as they were analysed from isogenic control midbrain organoids. It was observed in vivo that the insulin resistance affects human midbrain and accelerates PD phenotypes.

In order to further validate the relevance of these datasets for PD-T2DM commorbidity study, a dedicated analysis performed (see Figure B.22), involving the following steps.

1. Enrichment analysis of the differentially expressed genes (DEGs) to identify significant pathways and gain a better understanding of their role:

- Use Gene set enrichment analysis to represent the overlaps on the Substansia Nigra dataset [194] on the PD map to obtain common pathways.
- Use the MsigDB Hallmark 2020 database (www.gsea-msigdb.org) to identify common disease hallmarks across a wide range of diseases for better interpretation.
- Perform pathway enrichment analysis using the StringDB (https://string-db.org/) and EnrichNet tool [199] to create subnetworks and identify potential interactions.
- Use KEGG [2] and Reactome [200] to identify significant pathways and confirm their relevance with other results.
- Validate the enrichment results and brain tissue specificity using InnateDB [201], GTEx [202], and HPA [203].

2. A literature search for miRNA targets in PD and DM in patients by searching online databases and reviewing the identified articles for relevant information. This was followed by reviewing the identified articles and selecting those that report on validated miRNA targets commonly found in both PD and DM in patients. The results of the literature search were further filtered by the significant and relevant targets identified through the enrichment analysis.

3. Compare if any of the filtered common miRNAs targets have overlaps with miRNAs-target of PPMI dataset. The aim is to check the relevance between two diseases before the simulation.

The raw results from the enrichment analysis in different databases are available in the "Supplementary results directory" in GitLab repository [183]. The filtered outcome from the workflow is presented in Table 4.8

3.4.3 Stochastic Boolean model simulation

Simulations of the selected BMs were performed using pyMaBoSS, a python API for the MaBoSS software [169]. The pyMaBoSS framework is a tool for

probabilistic Boolean modelling and simulation of biological systems by applying discrete/continuous time Markov processes. It uses a Monte Carlo algorithm to simulate the evolution of the system over time, based on the initial conditions of the biomolecules and the rules governing their interactions. In this context, pyMaBoSS performs asynchronous updates using a random walk which is a simulation technique where a single biomolecule is selected and updated at each step and this process is repeated to create a sample of the reachable attractors. In pyMaBOSS, random asynchronous transitions are used to identify steady states and complex attractors from the specified initial states. The perturbation analysis includes only Knockouts and overexpressions.

The probabilities of the initial states are identified based on the effect size and statistical correlations of PPMI dataset. The effect size and statistical correlations of the dataset are used to assign probabilities to the initial states, which are then included in configuration files that can be used with PyMaBoSS.

Probabilistic Boolean simulation:

In pyMaBoSS, a biological system is represented as a model of interconnected Boolean variables, each representing the state of a biomolecule (e.g. present or absent, active or inactive). The interactions between the variables are defined by Boolean rules, which specify how the state of one variable can influence the state of another. The Monte Carlo algorithm works by randomly sampling the possible states of the system at each time step, based on the probabilities of each state given the current state of the system and the rules governing the interactions between variables. By simulating the system over many time steps, PyMaBoSS can estimate the probability of each state at each time point, allowing researchers to study the dynamics of the system and predict how it will evolve over time. [145].

Parametrisation using PPMI miRNA data Each biomolecule has a specific transition rate associated with its state. The probability of reaching a phenotype (observed output of a model) is calculated by simulating random walks over the probabilistic BMs. Thus, pyMaBoSS can be used to investigate the variation in phenotype probability as a result of a particular molecular alteration[172, 37, 165]. The study focuses on the model's outputs as well as some biomolecules that explain the variation in the progression of disease subtypes. These probabilities describe model states of disease subtypes in a series of iteration steps. The change points of the iteration steps are identified using the regression of multiple change points algorithm[204].

Parametrisation using T2DM transcriptomic profile The identified targets from previous steps were perturbed in order to study the effect of perturbation of T2DM on PD progression (Table 4.8). The perturbations were based on the expression levels of these biomolecules in datasets and involved point mutations, with knockout represented as ”Zero” and overexpression represented as ”One”. The downstream effects of these perturbations on PD progression were simulated. The probability of reaching a particular phenotype (observed output of the model) was calculated by simulating random walks on the probabilistic biomolecular models. These probabilities describe the different states of PD progression and the specific impact of perturbation of T2DM on PD progression.

Dynamic Time Warping:

Then, Dynamic Time Warping (DTW)[205] is an algorithm to measure similarity between two temporal sequences. The DTW algorithm works by dividing the time series into points and measuring the distance between corresponding points in different series. In this context, DTW was used to determine the similarities and differences between trajectories based on the calculated change points. The distance between points is calculated using the Euclidean distance, which is a measure of the straight-line distance between two points in a space. The DTW algorithm calculates the distance between every point in one series and the first point in a second series, and then finds the optimal path through these distances that minimizes the overall distance between the two series (Figure 4.4). A lower DTW score indicates a higher degree of similarity between the two series, while a higher DTW score indicates more differences between the series. Therefore, the DTW score can be used as a measure of the ”activity” or dynamics of a particular process or trajectory over time. A lower DTW score may indicate a higher level of activity or a more consistent progression of the condition, while a higher DTW score may indicate less activity or more variability in the progression of the process. However, the interpretation of the DTW score requires understanding of the specific characteristics of the disease conditions.

Pearson correlation:

Pearson correlation was used to measure the correlation between of DTW similarity values in pairs of subgroups (Table 4.17). The Pearson correlation coefficient is a statistical measure that was used to assess the strength and direction of the linear relationship between two variables [189]. In this context, a positive correlation indicates that as one variable increases, the other variable also tends to increase. A negative correlation indicates that as one variable increases, the other variable tends to decrease. The strength of the correlation is indicated by

the magnitude of the coefficient. A coefficient close to +1 or -1 indicates a strong relationship, while a coefficient close to 0 indicates a weak relationship. This information can be useful in the diagnostic and treatment planning process, as it can inform the selection of appropriate therapies and interventions for different disease conditions.

Chapter 4

Results

In this part, methods and algorithms discussed above were applied to study the dynamic behaviour of molecular mechanisms in Parkinson's disease (PD), and how they respond to perturbations. To achieve this, different formats of BMs were created - including SBML-qual- and analyzed with a range of tools. The accuracy of the BMs was assessed through the analysis of their structural and dynamic properties. The results of the assessment showed that the models have consistent structural and dynamic properties. This suggests that the selected models are reliable can be used in studying dynamics of PD.

The BMs were stratified by integrating omics data of patients in disease subgroups to study disease heterogeneity and specific responses to molecular perturbations. Molecular expression datasets were used to specify the parameters of simulations for selected subgroups. From this point, the modelling of disease subgroups can develop specific treatment strategies and propose new therapeutic interventions to arrest disease progression.

All results are available in the GitLab repository [1].

4.1 Model construction

The selection of PD map diagrams (Table 4.1)for downstream modelling and verification is based on several factors. First, the diagrams with phenotypic relevance are chosen, including those related to dopamine dyregulation, alpha-synuclein aggregation, neuoinflammation, oxidative stress, and mitochondrial dysfunction. These phenotypes help to understand the underlying mechanisms of PD. Next, the quality of the information in the diagrams is evaluated through the literature search to identify the recently validated and reported data. The selected diagrams are then assessed for the presence of sufficient information, including the

[1] `https://gitlab.lcsb.uni.lu/Ahmed7emdan/Boolean modelling of PD -Thesis`

biomolecules and their interactions. Additionally, the complexity of the models is considered, with the aim of finding a balance between the computational efficiency and the accuracy. The primary objective is to use the chosen diagrams to reproduce the previously validated and reported behaviours in vivo.

The selection of PD map diagrams for downstream modelling and stratification is based on pathway enrichment analysis of the Parkinson's Progression Markers Initiative (PPMI) dataset using the PD map. The significant pathways (Table 4.1 are dopamine transcription pathways, PI3k/AKT signalling, FOXO3 acivity, mTOR-MAPK signalling, and PRKN mitophagy. It is important to note that these pathways are complex and interconnected, and the consequences of their dysregulation vary depending on the specific disease subgroups. (All miRNA targets from the PPMI dataset are used in the enrichment analysis to ensure involvement of pathways covering all considered disease subtypes)

The selected pathways are exported from PD map in specific representations, such as bipartite and CellDesigner SBML, for further analysis and modelling. CellDesigner SBML diagrams of these patwhways are used to produce SBML-qual files using CaSQ. Then, BMs are constructed using bipartite and SBML-qual representations.

Pathways	Nodes	Edges	Type
Dopamine transcription	167	196	Cellular signalling
Wnt-PI3K/AKT signalling	391	436	Cellular signalling
FOXO3 activity	65	86	Cellular signalling
mTOR-MAPK signalling	59	83	Cellular signalling
PRKN mitophagy	54	72	Cellular signalling
PPARGC1A	67	109	Cellular signalling
TCA cycle	137	160	Metabolic

Table 4.1 This table presents a summary of the selected pathways with their nodes, and edges in several important cellular signaling and metabolic pathways. The pathways listed include dopamine transcription, Wnt-PI3K/AKT signaling, FOXO3 activity, mTOR-MAPK signaling, PRKN mitophagy, PPARGC1A, and the TCA cycle. The number of nodes and edges for each pathway is also provided

4.2 Model verification

Verification of a BM is important to ensure its accuracy and reliability. Verification of a BM involves examining both its structural and dynamic aspects. Structural verification involves evaluating the relationships and interactions between the model's components to ensure that they accurately reflect the underlying biological system. Dynamic verification involves evaluating the model's behavior over time and comparing it to the actual system to ensure that the model's predictions

are reliable. By verifying a BM in both of these ways, the model can be ensured to be a reliable and accurate representation of the biological system it is intended to model, which is essential for making informed and reliable predictions about that system. The verification was run for all the models selected previously to ensure that they are reliable and accurate representations of the biological systems.

4.2.1 Structural verification

The structural verification of the models was performed by using the topological network analysis and centrality measures as discussed in (Section 3.2).

General observations: Results of this analysis are in (Table 4.2). The table show that signalling pathways tended to have higher average degree, betweenness, and stress values, suggesting their importance in coordinating the system's response to stimuli. The "Dopamine transcription" pathway, on the other hand, had relatively low values in these measures, potentially indicating a less influential role in the overall system. Within the "TCA cycle" pathway, intermediates such as "2-oxoglutaric acid," "S-malate," "NADH," "ADP," and "acetyl-CoA" had high degree and stress values, indicating their importance in maintaining the flow of the cycle and regulating metabolism. In the "Wnt-PI3K/AKT" pathway, signaling molecules like "PI3K," "PDPK1," "RPS6KB1 phosphorylated," and "AKT1 phosphorylated" had high degree, betweenness, and stress values, suggesting their central roles in transmitting signals and regulating downstream targets.

In the "Dopamine transcription": The nodes with the highest in-degrees (TF NR4A2 complex, PITX3, and TF PITX3 complex) also had relatively high betweenness and stress values. This suggested that these nodes may play a central role in the pathway, possibly regulating the transcription of dopamine in response to various signals.

In the "FOXO3 activity": The nodes with the highest degree, betweenness, and stress values are all transcription factors or transcription factor complexes. This suggested that these factors may be key regulators of FOXO3 activity, possibly through the transcription of downstream target genes.

In the "mTOR": The pathway appeared to be highly interconnected, with many nodes having high degree, betweenness, and stress values. This suggested that mTOR activity may be tightly regulated by multiple signaling pathways and factors.

Pathway	Node ID	Degree			Betweenness	Stress
		Total	In	Out		
Dopamine transcription	TF NR4A2 complex	29	8	21	330	463
	PITX3	3	2	1	102	136
	TF PITX3 complex	13	2	11	66	84
	MIR133B rna	2	1	1	54	66
	RXRA	2	1	1	27	38
FOXO3 activity	TFFOXO3 complex nucleus	36	9	27	3673	9084
	TFCHOP:FOXO complex	4	2	2	1763	4932
	FOXO3 nucleus	3	2	1	1600	4416
	SIRT3	6	3	3	1482	3946
	BBC3 rna	3	2	1	896	2670
mTOR	AMPK complex neuron	10	7	3	2277	11064
	TSC1:TSC2 complex neuron	8	5	3	1465	7476
	STK11	6	5	1	1018	6050
	SESN2	5	2	3	788	2464
	nicotinamide	7	4	3	745	7518
PPARGC1A	PPARGC1A phosphorylated	13	4	9	322	650
	TF NRF1 complex	24	2	22	174	388
	PPARGC1A acetylated phosphorylated	4	3	1	94	188
	TF NRF2 complex	16	2	14	88	266
	TF YY1 complex	11	3	8	75	92
TCA cycle	2-oxoglutaricacid	25	19	6	451	761
	S-malate	12	7	5	195	490
	NADH	19	15	4	188	539
	ADP	15	8	7	178	393
	acetyl-CoA	10	6	4	152	384
	mTORC1 complex neuron	11	7	4	463	511
	AKT1 phosphorylated	7	5	2	440	500
Wnt-PI3K/AKT	PI3K	9	8	1	297	369
	PDPK1	2	1	1	260	326
	RPS6KB1 phosphorylated	5	3	2	189	201

Table 4.2 Common top five topological metrics in BMs and their source diagrams. The table summarises topological properties of nodes in the selected pathways, and "Node ID" indicate the specific nodes, "Degree" indicates a specific type of a node degree of in the BMs (total/incoming/outgoing connections), "Betweenness" describes node betweenness, while "Stress" describes the number of shortest paths that pass through a node

In the "PPARGC1A": The node with the highest betweenness value (TF NRF1 complex) also has a relatively high in-degree, but a low out-degree. This suggests that the NRF1 complex may be receiving signals from many other nodes in the pathway, but only influencing a few downstream targets.

In the "TCA cycle": The pathway had a number of nodes with high degree and stress values, but relatively low betweenness values. This suggests that these nodes are important for maintaining the flow of intermediates within the model, but may not be as central to the overall structure of the pathway.

The analysis found that all models passed the verification process and were appropriate for modelling. The results demonstrate the high quality and reliability of these models.

4.2.2 Sensitivity analysis

The dynamic verification of the models was performed by using the sensitivity analysis and attractor reachability as discussed in Section 3.3.

The results of the analysis show that all of the models are less sensitive to knockouts (in which a particular biomolecule is removed) than to overexpressions (in which a biomolecule is significantly increased). This suggests that the models are more robust to the removal of single biomolecules, but are more sensitive to significant increases in the levels of certain biomolecules. One possible reason for this result is that the removal of a biomolecule may not have a significant impact on the overall functioning of the system, while an increase in the levels of a biomolecule may disrupt the balance of the system and lead to more significant changes in the stable states. The models tend to be more sensitive to perturbations of complex biomolecules than simple biomolecules (Figures 4.1 and 4.2). Complex biomolecules, such as transcription factor complexes, had many interacting parts and play important regulatory roles in the models. As a result, disrupting the function of these molecules had significant sensitivity values for the overall functioning of the model.

In Figures 4.1 and 4.2, a higher identity-based and similarity based distances (see Section 3.3.3) indicated a larger difference between the original and perturbed attractors, which suggested a higher sensitivity of the models to perturbations. The data suggested that the sensitivity of the models to perturbations varies across the different pathways and biomolecule groups. For example, in (Figures 4.1 and 4.2), the model of the TCA cycle pathway was more sensitive to perturbations (knockouts) compared to the model of the PRKN pathway, as the distance values for the TCA cycle pathway were generally higher than those for the PRKN pathway. Similarly, the model of the PPARGC1A pathway was more sensitive to perturbations (knockouts) compared to the model of the FOXO3 pathway, as the

distance values for the PPARGC1A pathway were generally higher than those for the FOXO3 pathway.In (Figures 4.1 and 4.2), the CTNNB1, EIF4EBP1, IRS1, CTNNB1, and PDPK1, had ”Identity-based distance” values of 1. This indicates that the models were highly sensitive to perturbations (overexpression) of these groups in the corresponding pathways. The overexpression of these biomolecules could significantly alter the behavior of the model.

It’s worth noting that the ”Identity-based distance” and ”Similarity-based distance” values may not always be directly comparable across pathways and biomolecules, as these measures are calculated using different methods and may be affected by different factors. However, the observation of these pathways and biomolecules suggested that they may be key players in the dynamics of the model and could have significant downstream effects.

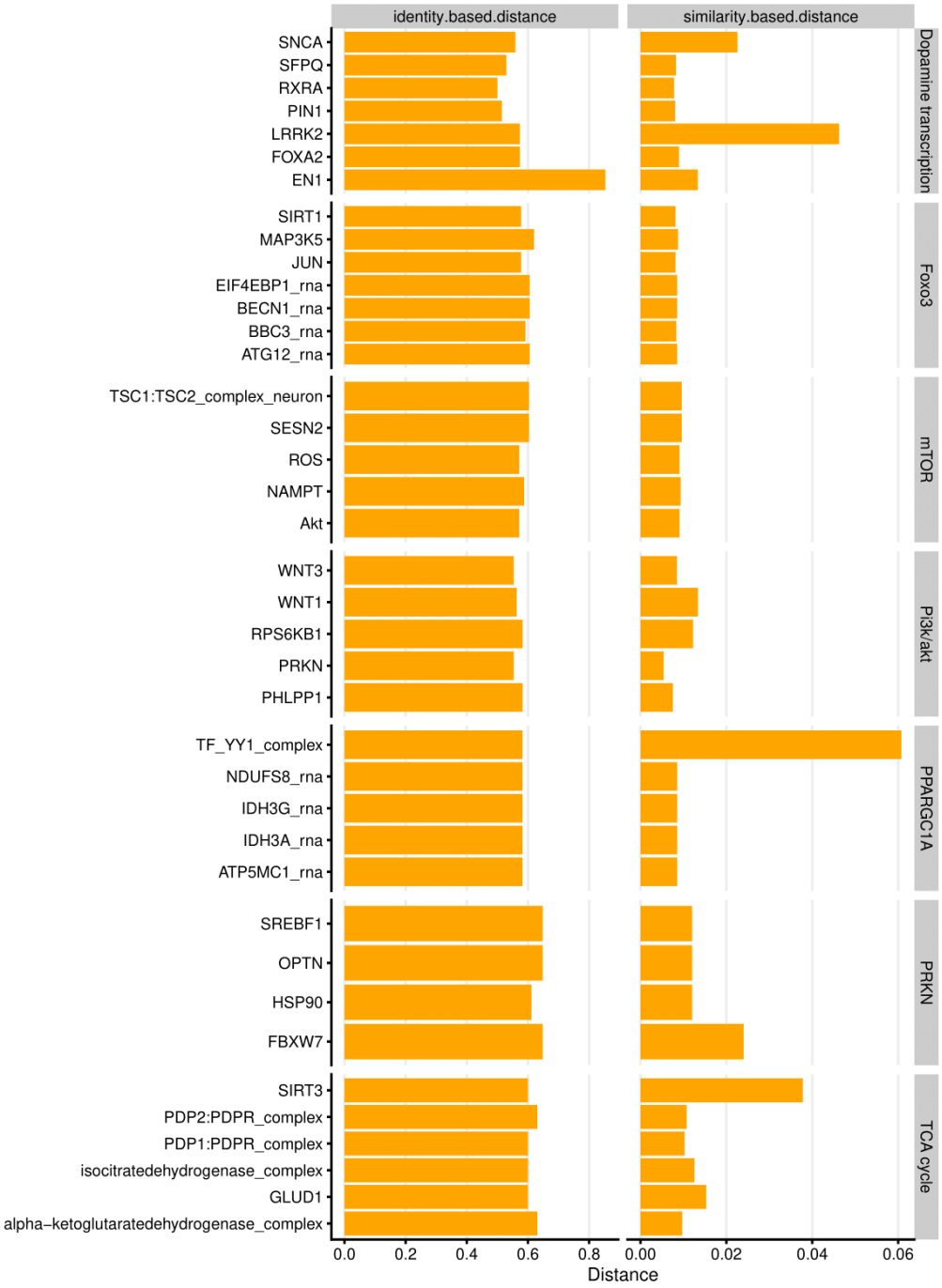


Figure 4.1 The top sensitivity metrics in BMs. The figure summaries the sensitivities of the BMs in response to molecular knockouts. The figure show the identity and similarity-based distances between the original and perturbed attractors

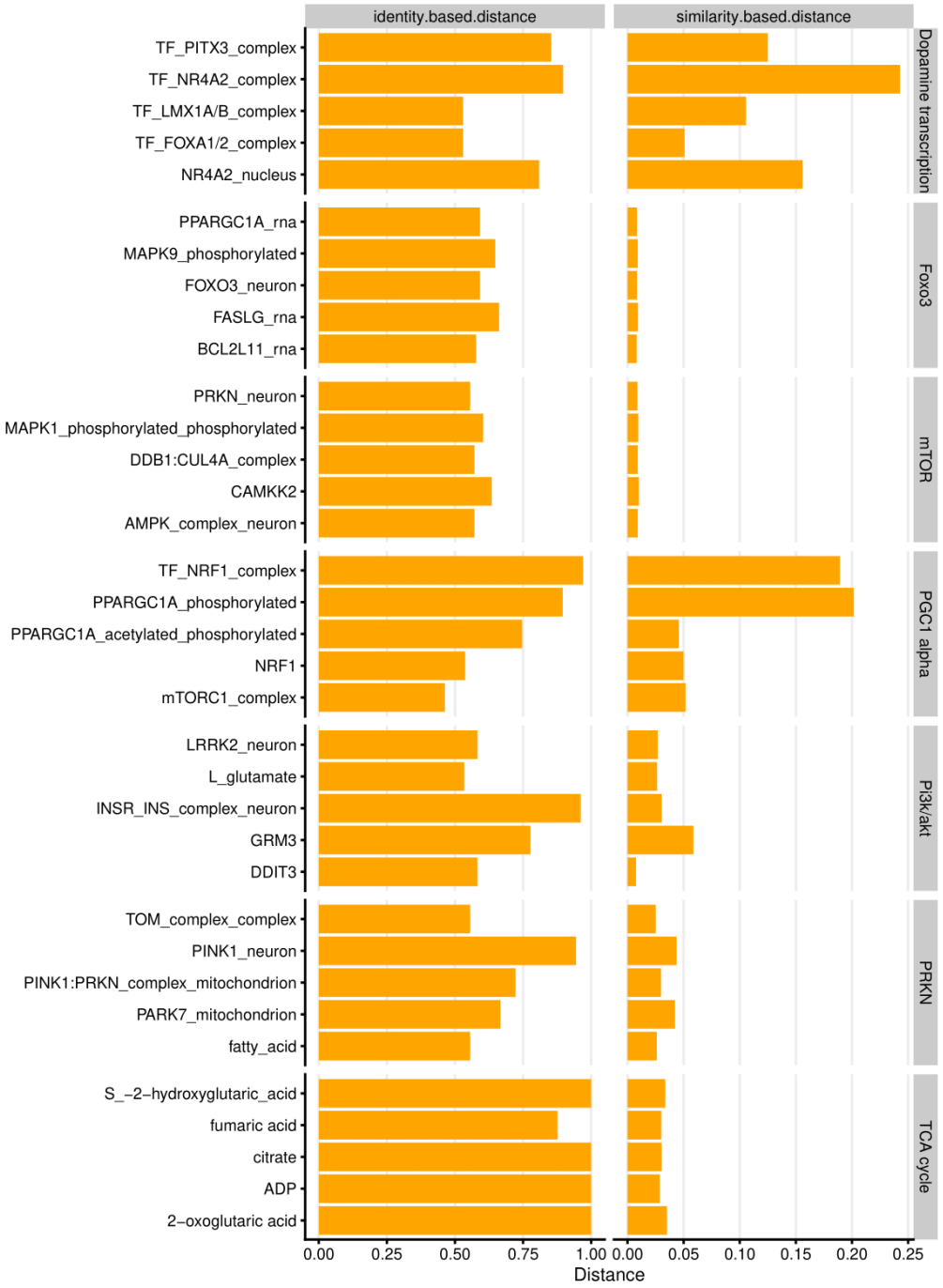


Figure 4.2 The top sensitivity metrics in BMs. The figure summaries the sensitivities of the BMs in response to molecular overexpressions. The figure show the identity and similarity-based distances between the original and perturbed attractors

4.2.3 Attractor reachability

Asynchronous and synchronous simulations were compared in the selected models to understand their characteristics and calculate their performance. The results of the attractor analysis showed that state trajectories converge to fixed or cyclic attractors under different updating schemes. The Table 4.3 summarises the comparison of the performance of four algorithms (HyTarjan, Heuristic, Decomp, and SAT) (see Section 3.3.2) in terms their reachability in pathways. The data show that in case of asynchronous solutions, HyTarjan outperforms the Heuristic approach. In turn, the comparison between Decomp vs SAT shows that there is not a single solution that consistently outperforms the other.

Next, the attractors produced by these methods were analysed in terms of their biological relevance (see Section 4.3 below). Both HyTarjan and Decomp solutions produced attractors that were not biologically relevant for all the considered pathways. In turn, Heuristic and SAT solver algorithms produced attractors that were viable for all the pathways except Wnt/PI3K pathway (see Section 4.3.2). The reason for this behaviour was an aggressive decomposition by the HyTarjan and Decomp algorithms, overly fragmenting the resulting models. Taken together, Heuristic and SAT solver algorithms were selected for downstream analysis.

The data showed that the SAT algorithm consistently outperformed the Decomp algorithm in terms of the time reachability, with the exception of TCA cycle. The SAT algorithm achieved greater time reduction in ER stress signalling pathway, reaching 87.32% improvement compared to the Decomp algorithm. Overall, the SAT solver algorithm improved the speed of attractor reachability by reducing the initial size of the states.

			Time (seconds)			
			Asynchronous		Synchronous	
Pathway	Edges	Targets	HyTarjan	Heuristic	Decomp	SAT
PGC1 alpha	109	PPARGC1A	3547	1789	173	96
		SIRT1	2587	1471	169	74
Dopamine transcription	167	NR4A2	2981	1460	147	54
Wnt/PI3K-AKT	391	Wnt/PI3K	1135	3961	403	256
ER stress signaling	53	DDIT3	971	1855	67	11
		AKDHC	2066	1123	84	110
TCA cycle	137	Oxoglutarate	2122	1140	84	110
		IDH	2153	1151	84	110
		SIRT3	2130	1140	84	110

Table 4.3 Attractor reachability speed in different algorithms. The table shows the duration of attracor reachabiliy for asynchronous and synchronous systems in the selected pathways, using the methods HyTarjan, Heuristic, Decomp, and SAT. The scales includes the node numbers, and Targets indicate the perturbed molecules

4.3 Model validation

Validation is the process of comparing a model's output to real-world data to confirm its reliability. To validate a model, simulations are performed and the model's ability to reproduce known perturbations is evaluated. All the selected BMs were validated following the steps indicated in Section 3.2. The following results suggested that they are reliable indicators of biological processes, as they are consistent with the available data. However, there are two cases in Wnt-PI3K/AKT signaling pathway where the model did not match the published literature, which led to its correction.

For each pathway, the selected targets were perturbed following the literature-suggested scenarios, and the changes in corresponding outputs were analysed. This was done by:

- a) Running a simulation in CellCollective [7] and analysing the attractors visually
- b) Performing an exhaustive attractor search with pyMaBoSS (available in the GitLab repository [183]) and comparing computationally the obtained attractors with DEGs reported in corrseponding published datasets.

Details of validation scenarios for selected pathways, their results and interpretation are discussed below.

4.3.1 TCA cycle model: Validation using literature and experimental data

The TCA cycle is a critical metabolic pathway that occurs in the mitochondria of cells. This pathway is impaired in PD, and this impairment may play a role in the development of the disease [206]. The TCA cycle is a good example to validate the BM approach because it is a well-studied and well-understood biological process, with a wealth of literature and data available for comparison [207]. Additionally, the TCA cycle involves the conversion of a number of different metabolites, which can be measured and tracked using techniques such as fluxomics. This allows to evaluate the accuracy of a BM by comparing the predicted metabolite levels to those observed experimentally. Given that the TCA cycle is a central metabolic pathway that is involved in many different cellular processes, a BM of the TCA cycle could have broad implications and applications.

The performance and accuracy of the TCA cycle BM are validated through comparison with literature and fluxomics data. In order to run the simulation, the knockouts of isocitrate dehydrogenase, alpha-ketoglutarate dehydrogenase, and the pyruvate dehydrogenase complex are considered as initial state parameters

based on literature. Using temporal-fluxomics data [207], the activity levels of alpha-ketoglutarate dehydrogenase and the nucleotide triphosphates GTP and GDP on ATP levels are simulated using CellCollective. After running the simulation with these parameters, the results are compared with experimental data as discussed below. The validated results based on literature show that the simulated activation of acetyl CoA, NADH, and PDKs through increases the phosphorylation reaction, which in turn reduces the activity of the pyruvate dehydrogenase complex (PDC). This supports the hypothesis that PDC deficiency, a potential therapeutic target for age-related diseases, results from the activation of the phosphorylation reaction involving these products [208]. The PDC deficiency, which is caused by KGDHC knockout, has a significant effects on the production of succinic semialdehyde [209]. This deficiency leads to a decrease in the levels of succinic acid and succinyl CoA, resulting in a decline in ATP and GTP in mitochondria [210, 211]. Isocitrate dehydrogenase knockout decreases the amount of ATP and inhibit the oxidative decarboxylation catalysis of isocitrate into alpha-ketoglutarate, leading to mitochondrial dysfunction and dopaminergic neurotoxicity [212]. In PD, the key regulators such as oxoglutaric acid, glutamate hydrogenase 1 (GLUD), and ATP levels are deregulated in response to SIRT3 knockout, which directly impacts mitochondrial function [206]. These findings are summarized in the table (Table 4.6).

The cellular metabolites oscillate during the cell cycle, adapting the changes in the cell. The effects of alpha KGDH, and GTP, GDP are simulated on ATP level.

		Inputs				Output ATP level	
Time (h)	Cond.	α-KG acid	GTP	GDP	Phosphate	**Simulated**	**Real**
9	C1	0.84	0.48	0.50	0.45	1	0.87
12	C2	0.79	0.77	0.66	0.67	0.86	0.77
15	C3	0.77	0.98	0.79	0.93	1	1
18	C4	0.84	0.83	0.83	0.71	1	0.95
21	C5	1	0.56	0.65	0.73	1	0.90
24	C6	0.88	0.50	0.64	0.72	0.96	0.84
27	C7	0.93	0.66	0.84	0.80	0.96	0.87
30	C8	0.83	0.93	1	0.78	1	0.89
33	C9	0.83	1	0.71	1	1	0.89
36	C10	0.80	0.79	0.61	0.93	0.9	0.81
39	C11	0.81	0.77	0.60	0.83	1	0.97
42	C12	0.80	0.73	0.58	0.73	0.95	0.85
45	C13	0.815	0.49	0.46	0.65	0.88	0.76

Table 4.4 This table presents data on the ATP activity in the TCA (tricarboxylic acid) cycle with fluxomics integration. The data includes the time in hours, the condition (labeled as "Cond."), and the levels of various inputs (alpha-ketoglutaric acid, GTP, GDP, and phosphate). The table also includes both simulated and real measurements of ATP levels.

The simulated ATP activity levels are inline with the measured concentration in synchronised HeLa cells every 3 hours post release from growth arrest (Table 4.4).

4.3.2 Signalling pathways: Validation using literature and experimental data

To further validate the capability of BMs to simulate biological processes, a group of signaling pathways were chosen as benchmarks. They were selected based on the disease relevance and available experimental data to validate against the predictions.

Dopamine transcription pathway

One of the validated pathways in this study was the dopamine transcription pathway, which is known to be deregulated in Parkinson's disease (PD)(Figure A.2) [213].

The key elements considered in this validation are: NR4A2 and SIRT1. NR4A2 and SIRT1 are key elements considered in this validation because they are proteins that play important roles in the development and maintenance of neurotransmitters and various cellular processes [213, 214, 215].

In the simulation scenario, NR4A2 and SIRT1 were perturbed and the effects on dopamine release and mitochondrial biosynthesis were observed as activity levels using the CellCollective. These results, which were confirmed through an exhaustive search attractor analysis (available in the GitLab repository [183]), show the following behavior that is in line with the following literature (Table 4.6):

- The production of brain-derived neurotrophic factor (BDNF), which is important for the survival and growth of neurons, was affected by the NR4A2 knockout (Figure 3.5). This result was in line with previous research indicating that NR4A2 is involved in the stimulation of BDNF production in response to the neurotransmitter N-methyl-D-aspartate (NMDA) [213].

- In addition, other molecules such as GCH1, TH, DDC, SLC18A2, SLC6A3, and DRD2, which have significant effects on the development and maintenance of neurotransmitters through various targets, were affected by the NR4A2 knockout, as stated in previous researches [216, 217, 218].

- Activation of SIRT1 was also observed to improve the biogenesis of mitochondria, which are important for energy production in cells, according to previous research [214].

Wnt-PI3K/AKT

The PI3K/AKT pathway was one of the validated pathways in this study and is known to be affected by mutations in PD [219, 220, 221].

The key elements considered in this validation are: LRRK2 G2019S mutation, Wnt, DDIT3, GSK3B, TFEB, PHLPP, RPS6KB1, AND 4EBP2. These molecules are involved in the Wnt-PI3K/AKT pathway and plays a role in the development and progression of PD [220, 125, 222, 223].

In the simulation scenario, these biomolecules were perturbed and the effects on autophagy and neuron survival were observed as activity levels using the CellCollective.These results, which were confirmed through an exhaustive search attractor analysis (available in the GitLab repository [183]), show the following behavior that is in line with the following literature (Table 4.6, (Table 4.5)):

- Overexpression of the LRRK2 G2019S mutant enhanced the autophagy process [220].

- This effect appeared to be mediated by the inhibition of mTORC1/2, proteins that regulate autophagy [125, 222]. These observations supported the idea that amino acid sensing, mTORC1 signaling, and autophagy were closely interconnected and suggested that modulating these pathways may be useful for treating PD [221].

- Activating both 4EBP2 and TFEB was shown to increase autophagy activity more than activating them separately [224, 225, 226]. Specifically, activating both of these proteins in combination leaded to a 10.7% and 13.6% increase in autophagy compared to activating them individually.

- In addition, the inhibition of the proteins RPS6KB1 and PHLPP in combination with the activation of TFEB significantly decreased neuronal death and the active state of autophagy [227]. This combination resulted in a 96.3% decrease in neuronal death.

- Activating the Wnt protein and inhibiting the protein GSK3B both increase autophagy [228, 229, 65].Further, the combination of modulating Wnt signaling and GSK3B activity may improve our understanding of therapeutic protocols for neurological diseases by promoting neurogenesis and autophagy [39]. These findings suggested that targeting these proteins could be a promising approach for developing treatments for neurological disorders (Table 4.5).

Inputs							Outputs	
4EBP2	TFEB	gsk3b	Wnt	PHLPP	RPS6KB1	LRKK2	Autophagy	Neuron death
1	–	–	–	–	–	–	0.3983	0.9557
–	1	–	–	–	–	–	0.3691	0.9934
–	–	0	–	–	–	–	0.6918	0.9592
1	1	–	–	–	–	–	0.5055	0.8083
–	–	0	1	–	–	–	0.9928	0
–	1		–	0	0	–	0	0.0364
1	1	0	–	0	0	1	0.9926	0.0364

Table 4.5 The table shows the relationship between molecular target interventions and the probabilities of autophagy and neuron death. The data includes the levels of seven different inputs, which are known to play a role in these processes. The table presents the resulting probability of autophagy and neuron death for each combination of input levels

FOXO3 activity

One of the validated pathways in this study was the FOXO3 activity pathway, which is known to be dysregulated in PD [230].

The key element considered in this validation is: FOXO3. FOXO3 is a key element considered in this validation because it plays important roles in autophagy, cell cycle progression, apoptosis, stress resistance in PD [230, 231, 232].

In the simulation scenario, FOXO3 was perturbed and the effects on autophagy and RNA mediated bioomolecules were observed as activity levels using the CellCollective. These results, which were confirmed through an exhaustive search attractor analysis (available in the GitLab repository [183]), show the following behavior that is in line with the following literature (Table 4.6):

- FOXO3 biomolecule activation increased autophagy in mitochondria (Table 4.6) [232].
- FOXO3 activation also activated different molecules involved in RNA mediated mechanisms, including BECN1, GABARAPL1, MAP1LC3A, BNIP3, ATG12, and MUL1. These biomolecules are known to be important regulators of autophagy [233].

4.3.3 Cross-pathway comparison

In order to validate the relevance of BMs for studying mechanisms of PD, one metabolic and three signalling pathways were analysed. Literature-driven scenarios were used to validate the pathways, as shown in the Table 4.6. The simulated behavior of these pathways matched the expected behavior according to published literature. However, the Wnt/PI3K pathway showed some discrepancies.

Namely, the simulated behavior of LRRK2 is not matching the expected behavior in some researches [234, 235]. According to the "LRRK2: Autophagy and Lysosomal Activity" review article,[219]. There could be several reasons why there are discrepancies between the simulated behavior of LRRK2 and its expected behavior. Some possible reasons could include:

- Differences in the experimental conditions or protocols used to study LRRK2, which could lead to varying results.
- The use of different LRRK2 models or cellular systems in different studies, which could affect the results.
- The complexity of the autophagy pathway and the many factors that can influence it, making it difficult to accurately predict the behavior of LRRK2 in all cases.
- Limited understanding of the exact mechanisms by which LRRK2 regulates the autophagy pathway, which could lead to discrepancies in the simulations.

	Dimension					
Pathway	Nodes	Edges	Target node	State	Simulated behavior	Expected behavior
TCA cycle	137	160	AKDHC	OFF	acetyl coA-ATP-NADH	Match [210]
			Oxoglutarate	OFF	acetyl coA-ATP-NADH	Match [212]
			IDH	OFF	Acetyl coA-ATP-oxoglutarate	Match [212]
			SIRT3	OFF	Acetyl coA-ATP-Iron	Match [212]
Dopamine transcription	167	196	NR4A2	OFF	Dopamine release	Match [236]
			SIRT1	ON	Mitochondrial biosynthesis	Match[214]
Wnt/PI3K	391	436	LRRK2	ON	Autophagy activation	Match[220], Mismatch [234]
			Wnt	ON	Increase auto-phagy	Match [58]
			DDIT3	ON	increase BCL2L11/BBC3i	Match [223]
			GSK3B	OFF	Autophagy activation	Match [229]
			TFEB	ON	Autophagy activation	Match [237]
			PHLPP	OFF	Autophagy dysregulation	Match [227]
			RPS6KB1	OFF	Autophagy dysregulation	Match [227]
			4EBP2	ON	Autophagy activation	Match [238]
FOXO3 activity	65	86	FOXO3	ON	Autophagy activation	Match [232]
					BNIP3 activation	Match [232]

Table 4.6 The table compares the simulated behavior of several Boolean models to expected behavior based on published literature. The table includes information on the pathways, the number of nodes and edges in each network, the target node, the state of the target node (ON or OFF), and the simulated and expected behavior for each pathway.

As a result, corrective measures were taken to address these discrepancies. These measures include modifications of Boolean function to better represent the interactions between LRRK2, ARFGAP1 and autophagy [239]. These results suggest

that the simulated behaviour of the pathways is largely consistent with published literature.

4.4 Modelling-based patient stratification using omics data

This part of the work focuses on modelling-based stratification of PD patients based on cohort-specific miRNA expression data. For this purpose, the PPMI dataset was used that includes microRNAs from blood samples in disease subtypes (LONI archive- www.ppmi-info.org/data). In the PPMI dataset, the following patient subgroups were considered:

1. **Prodromal** - 223 individuals - which does not develop severe symptoms but shows significant positive dopamine transporter (DAT) SPECT
2. **SWEDD** (Scan without evidence of dopaminergic deficit) - 187 individuals - clinically diagnosed with normal DAT SPECT
3. **Parkinsonism** - 123 individuals - includes cases of atypical PD, which is characterized by idiopathic symptoms similar to those seen in typical PD.

Moreover, the PD commorbidity with Type 2 Diabetes Mellitus (T2DM) was studied. For this purpose, the T2DM transcriptomic dataset (see Section 3.4.2) was used, describing the effect of PD-specific mutations on cell cultures of iPS-induced dopaminergic neurons and organoids.

These two datasets were used to parameterise and run selected models acquired form the PD map. Results discussed below describe analysis of the above-mentioned data, and simulation results for selected cohort subgroups, with and without the burden of T2DM commorbidity.

4.4.1 PD omics data analysis

miRNA data analysis for PD subgroups

The miRNA data was analyzed by calculating differential expression of miRNAs statistical correlations (see Section 3.4.1. Standardized effect sizes were then calculated and transformed into probabilities using the Common Language Effect Size method (see Section 3.4.3). Based on manually curated miRNA databases and the PD map, miRNA targets were identified and filtered.

The validated miRNA targets from manually curated miRNA databases are identified (see Section 3.4.1). The targets were filtered based on the substantia nigra dataset from the PD map (see Section 3.4.1) and compared with those reported

in published literature and datasets. It was found that most of the significant miRNAs were downregulated in PD and were involved in mitochondrial dysfunction. The miRNAs that did not match were discarded from the analysis (Table 4.7).

The majority of the identified miRNAs are dysregulated in all cohorts (see Table 4.7). The effect size of the filtered miRNAs differs between cohorts (see Figure 4.3).

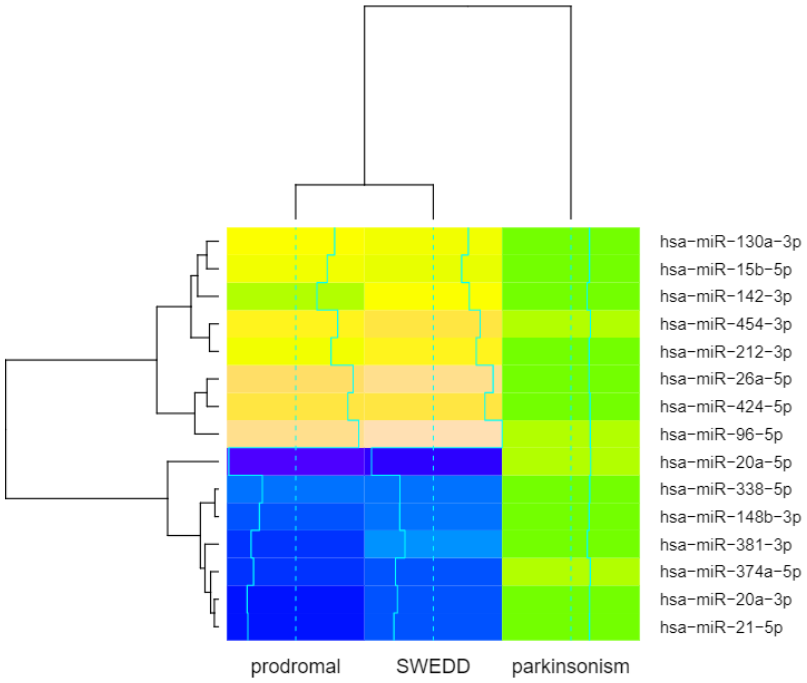


Figure 4.3 The figure shows an example of 15 miRNAs that are common across the conditions with different expressions

RNAseq data analysis for T2DM comorbidity

To identify potential connections between PD and T2DM, transcriptomic data from PINK1 and GBA mutations in T2DM was analysed Section 3.4.2. and comparing them to a substantia nigra dataset on PD. The following results were obtained:

1. Differentially expressed genes in two datasets describing the PINK1 and GBA N307S mutations in T2DM were identified and common overlaps with the genes in the substantia nigra dataset on the PD Map were determined [183].

2. Validated common miRNA-target pairs that was reported in the literature as being involved in both PD and T2DM were identified (Table 4.8) through

enrichment analysis of the differentially expressed genes (DEGs) and a literature search (see Chapter 3).

3. The downstream effects of T2DM on PD progression were simulated by perturbing the overlapped targets based on their expression levels in the datasets as point perturbations (knockouts and overexpressions).

The raw results from the enrichment analysis in different databases are available in the ”Supplementary results directory” in the GitLab repository, and the filtered outcome from the workflow is presented in the manuscript (Table 4.8).

	miRNA	Regulation	Sample	Method	Ref.
Matched expressions	hsa-miR-96-5p	Up	Peripheral blood	RT-qPCR (TaqMan)	[240]
	hsa-miR-26a-5p	Up	Peripheral blood	RT-qPCR (SYBR Green)	[241]
	hsa-miR-424-5p	Up	Peripheral blood	Microarray	[241]
	hsa-miR-9-3p	Up	Peripheral blood	RT-qPCR (TaqMan)	[241]
	hsa-miR-454-3p	Up	Peripheral blood	RT-qPCR (TaqMan)	[242]
	hsa-miR-15b-5p	Up	Peripheral blood	RT-qPCR (TaqMan)	[242]
	hsa-miR-671-5p	Up	Peripheral blood	Microarray	[243]
	hsa-miR-93-5p	Up	Prefrontal cortex	Illumina's HiSeq 2000	[243]
	hsa-miR-195-5p	Up	Peripheral blood	RT-qPCR (TaqMan)	[244]
	hsa-miR-20a-5p	Up	Peripheral blood	Microarray	[244]
	hsa-miR-16-5p	Up	Prefrontal cortex	Illumina's HiSeq 2000	[244]
	hsa-miR-132-3p	Up	Peripheral blood	RT-qPCR (TaqMan)	[245]
	hsa-miR-196b-5p	Down	Peripheral blood	RT-qPCR (TaqMan)	[245]
	hsa-miR-92b-3p	Down	Peripheral blood	Microarray	[245]
	hsa-miR-19a-3p	Down	Mid-brain	Microarray	[245]
	hsa-miR-19a-3p	Down	Peripheral blood	Microarray	[245]
	hsa-miR-92a-3p	Down	Peripheral blood	RT-qPCR (TaqMan)	[245]
	hsa-miR-133b	Down	Mid-brain	RT-qPCR (TaqMan)	[245]
	hsa-miR-15b-5p	Down	Peripheral blood	RT-qPCR (TaqMan)	[245]
	hsa-miR-7-5p	Down	Peripheral blood	RT-qPCR (TaqMan)	[240]
	hsa-miR-15a-5p	Down	Peripheral blood	Microarray	[246]
	hsa-miR-19b-3p	Down	Peripheral blood	RT-qPCR (TaqMan)	[247]
	hsa-miR-139-5p	Down	Peripheral blood	RT-qPCR (TaqMan)	[247]
	hsa-miR-450b-5p	Down	Peripheral blood	RT-qPCR (TaqMan)	[247]
	hsa-miR-212-3p	Down	Prefrontal cortex	Illumina's HiSeq 2000	[247]
	hsa-miR-22-3p	Down	Peripheral blood	RT-qPCR (TaqMan)	[247]
	hsa-miR-26a-5p	Down	Peripheral blood	RT-qPCR (TaqMan)	[247]
	hsa-miR-16-2-3p	Down	Prefrontal cortex	Illumina's HiSeq 2000	[247]
	hsa-miR-16-2-3p	Down	Peripheral blood	RT-qPCR (TaqMan)	[248]
	hsa-miR-30b-5p	Down	Peripheral blood	RT-qPCR (TaqMan)	[248]
	hsa-miR-144-3p	Down	Prefrontal cortex	Illumina's HiSeq 2000	[249]
	hsa-miR-323a-3p	Down	Peripheral blood	RT-qPCR (TaqMan)	[250]
	hsa-miR-495-3p	Down	Peripheral blood	RT-qPCR (TaqMan)	[250]
	hsa-miR-148b-3p	Down	Peripheral blood	RT-qPCR (TaqMan)	[250]
	hsa-miR-374a-5p	Down	Peripheral blood	RT-qPCR (TaqMan)	[250]
	hsa-miR-199b-3p	Down	Peripheral blood	RT-qPCR (TaqMan)	[250]
	hsa-miR-374b-3p	Down	Peripheral blood	RT-qPCR (TaqMan)	[250]
	hsa-miR-20a-5p	Down	Peripheral blood	Microarray	[250]
Mismatched expressions	hsa-miR-199b-3p	Up	Peripheral blood	Microarray	[250]
	hsa-miR-196b-5p	Up	Peripheral blood	RT-qPCR (TaqMan)	[251]
	hsa-miR-221-3p	Up	Peripheral blood	RT-qPCR (TaqMan)	[251]
	hsa-miR-103a-3p	Up	Peripheral blood	RT-qPCR (TaqMan)	[251]
	hsa-miR-320b	Up	Prefrontal cortex	Illumina's HiSeq 2000	[251]
	hsa-miR-30c-5p	Down	Peripheral blood	RT-qPCR (TaqMan)	[251, 240]
	hsa-miR-30a-5p	Down	Peripheral blood	RT-qPCR (SYBR Green)	[251, 248]
	hsa-miR-181c-5p	Down	Peripheral blood	RT-qPCR (TaqMan)	[252]
	hsa-miR-338-5p	Down	Prefrontal cortex	Illumina's HiSeq 2000	[253]
	hsa-miR-148b-3p	Down	Prefrontal cortex	Illumina's HiSeq 2000	[254]
	hsa-miR-21-5p	Down	Peripheral blood	Microarray	[241, 254]

Table 4.7 The table indicate the matched (top) and mismatched (bottom) expressions between the filtered PPMI-miRNAs and the reported miRNAs in literature and datasets. The table includes the miRNA name, the direction of regulation (up or down), the sample type, the method used for measurement, and the reference for the data.

		References	
miRNA	Targets	PD	T2DM
hsa-miR-423-3p	CDKN1A	[255]	[256]
hsa-miR-132-3p	MAPK1	[255]	[257, 258]
hsa-mir-15a-5p	RET, PHLPP1	[255]	[259, 260]
hsa-mir-29c-3p	PTEN	[261]	[262]
hsa-mir-29a-3p	IGF1	[261, 263]	[264, 262]
hsa-mir-20a-5p	PTEN, E2F1	[255, 265]	[266, 267]
hsa-mir-22-3p	PTEN	[268]	[269]
hsa-mir-26b-5p	IGF1R, PTEN	[270]	[271]
hsa-mir-143-3p	IGF1R, AKT1	[255]	[272]
hsa-mir-145-5p	IGF1R, IRS1, EIF4E, RPS6KB1	[273]	[274]
hsa-mir-133b	IGF1R, AKT1	[275]	[276]
hsa-mir-34a-5p	E2F1	[277]	[278]
hsa-mir-182-5p	PTEN, GSK3B	[279]	[280]
hsa-mir-148a-3p	IRS1	[270]	[281]
hsa-mir-7-5p	SNCA, IGF1R, RS1	[270, 255]	[282]
hsa-mir-195-5p	RET, INSR	[270]	[283]
hsa-mir-218-5p	RET	[284]	[285]
hsa-mir-200c-3p	ROCK2	[265]	[286]
hsa-miR-125b-2-3p	IGF1R	[287]	[288]
hsa-mir-18a-5p	PTEN, PHLPP1	[265]	[289]
hsa-miR-17-5p	PHLPP1, PTEN,E2F1	[284]	[290]
hsa-mir-96-5p	GSK3B	[291]	[292]
hsa-mir-21-5p	PTEN, E2F1	[293]	[294]
hsa-mir-200a-3p	PTEN, MAPK14	[295]	[296]
hsa-miR-200b-3p	PHLPP1, ROCK2	[295]	[296]
hsa-miR-200c-3p	ROCK2	[295]	[296]
hsa-mir-103a-3p	PTEN	[255]	[296]
hsa-mir-10a-5p	PTEN	[297]	[298]
hsa-mir-153-3p	SNCA, PTEN	[297]	[299]
hsa-mir-19b-3p	PTEN	[297]	[300]
hsa-mir-155-3p	PTEN	[263]	[301]
hsa-mir-26a-5p	GSK3B, PRKCD, PTEN	[263]	[302]
hsa-miR-26b-5p	IGF1R, PTEN	[263]	[302]
hsa-let-7a-5p	E2F1	[263]	[303]
hsa-mir-23a-3p	PTEN	[268]	[304]
hsa-mir-100-5p	AKT1, IGF1R	[305, 306]	[307]
hsa-mir-92a-3p	PHLPP1, PTEN	[305]	[308]

Table 4.8 The table indicates the common miRNA expression and target regulation in PD and T2DM. The table lists miRNA, targeted genes, and references for studies that have common altered expression of the miRNA in PD or T2DM.

4.5 Cohort specific Boolean simulations

Using the PD cohort-specific miRNAs and their identified targets, enrichment analysis was performed on the PD map using the MINERVA GSEA Plugin (3) to identify significant pathways affected by these miRNAs targets. These pathways include the dopamine transcription pathway, the Wnt-PI3KAKT pathway, the FOXO3 activity pathway, the mTOR-MAPK signalling pathway, and the PRKN mitophagy pathway (Table 4.9). These enriched pathways were then transformed into BMs and parameterized based on the effect sizes calculations (see Section 3.4.3 and Table 4.10).

Model	p value	Simulation inputs	Simulation outputs
Dopamine transcription	1.42E-14	ADCYAP1, BDNF, EN1, FOXO1, GCH1, PBX1, PRKAA2, RXRA, SESN3, SFPQ, RGS6, NRF1, MAP1B, LMX1A, FOXO3	Mitochondrial biogenesis Dopamine metabolism Neuron survival
Wnt-PI3KAKT	2.83E-24	AKT1, E2F1, EIF2AK3, GSK3B, IGF1, IGF1R, IRS1, MAPK1, NEDD4, PRKCD, PTEN, TFDP1, AGO2, EIF4E, IDE, PHLPP1, PPP2CA, PPP2CB, ROCK2, RPS6KB1	TFEB phosphorylated Insulin resistance TFEB SNCA complex TFEB complex Neuron death
FOXO3 activity	1.30E-20	AKT1, BCL2L11, CEBPB, FASLG, GABARAPL1, MAP3K5, MFN2, PPARGC1A,RICTOR, SESN3, SIRT1, SOD2, TXNIP, ATG12, BNIP3, FOXO3, HSPD1, JUN	Response to oxidative stress Fission Fusion Autophagy Mitochondrial biogenesis Apoptosis
mTOR-MAPK signalling	1.39E-22	AKT1, DEPTOR, GSK3B, MAPK1, MTOR, PRKAA1, PRKAA2,RHEB, RICTOR, RRAGD, SIRT1, TSC1, UBE2V1, CAMKK2, DDIT4, DEPDC5, DEPTOR, GSK3B,MAPK1, MAPKAP1, MTOR, PARP1,PHLPP1, PRKAA1, RHEB, RICTOR, RPS6KB1, SIRT1, TSC1, UBE2V1	Glycolysis RHEB lysosome AKT Catabolism Autophagy Mitochondrial biogenesis
PRKN	4.47E-05	ATXN3, BAG4, FBXW7, GABARAPL1, TIMM17A, ULK1, VPS13C	Apoptosis Mitophagy PRKN ubiquitinated PINK1 accumulation

Table 4.9 This table presents the Boolean models, with the inputs and outputs for each model listed. The pathways included are the Dopamine transcription pathway, the Wnt-PI3KAKT pathway, the FOXO3 activity pathway, and the mTOR-MAPK signalling pathway. The inputs for each pathway consist of various biomolecules, while the outputs represent various cellular processes or biomolecules that are influenced by the inputs.

4.5.1 Model parameterisation using cohort data

The probabilities of the initial states of BMs are determined based on the effect size and statistical correlations found in PPMI dataset, specifically for each subgroup (see Section 3.4.1). These probabilities are then included in configuration files and used with pyMaBoSS, a software tool that simulates the dynamics of BMs and investigates the probability of reaching certain conditions (levels of defined outputs).

Such parameterised models are then simulated with pyMaBoSS to study the effects of molecular alterations on the probability of reaching different conditions, as well as identify key biomolecules that may be involved in the progression of disease subtypes [172, 37, 165]

An example of such a parameterisation for the PRKN mitophagy pathway is discussed here. Table 4.10 illustrates calculated miRNA effect sizes for the SWEDD, prodromal, and parkinsonism, together with their targets. A corresponding BM, constructed based on a PD map diagram (see Figure A.6) is parameterised by setting the values of the corresponding targets (biomolecules in the model) to the corresponding probabilities. Such a parameterised model, run with pyMaBoSS provides readouts for the three phenotypes as illustrated by the corresponding lines (see Figure 4.4).

In the resulting simulation graphs (Figure 4.4) the change points in the iterations are identified using the regression of multiple change points algorithm [204].Two change points were chosen, dividing the graph into three sections,

Cohort	Target.Score	miRNA	Gene.ID	Gene.Symbol	Transcript.Accession	CL-effectsize
SWEDD	95	hsa-miR-96-5p	4287	ATXN3	NM_001164776	0.966761
	97	hsa-miR-26a-5p	9530	BAG4	NM_001204878	0.712208
	97	hsa-miR-424-5p	55294	FBXW7	NM_033632	0.852718
	95	hsa-miR-15b-5p	23710	GABARAPL1	NM_031412	0.92637
	95	hsa-miR-3121-3p	10440	TIMM17A	NM_006335	0.981296
	96	hsa-miR-26a-5p	8408	ULK1	NM_003565	0.712208
Prodromal	95	hsa-miR-96-5p	4287	ATXN3	NM_001164776	0.956198
	97	hsa-miR-26a-5p	9530	BAG4	NM_001204878	0.951438
	97	hsa-miR-15b-5p	55294	FBXW7	NM_033632	0.960511
	95	hsa-miR-15b-5p	23710	GABARAPL1	NM_031412	0.960511
	95	hsa-miR-3121-3p	10440	TIMM17A	NM_006335	0.574494
	96	hsa-miR-26a-5p	8408	ULK1	NM_003565	0.951438
Parkinsonism	96	hsa-miR-1271-5p	4287	ATXN3	NM_001164776	0.976557
	97	hsa-miR-26b-5p	9530	BAG4	NM_001204878	0.896051
	100	hsa-miR-32-5p	55294	FBXW7	NM_033632	0.914474
	96	hsa-miR-195-5p	23710	GABARAPL1	NM_031412	0.938137
	99	hsa-miR-421	9868	TOMM70	NM_014820	0.942465
	96	hsa-miR-26b-5p	8408	ULK1	NM_003565	0.896051
	98	hsa-miR-223-5p	54832	VPS13C	NM_018080	0.964472

Table 4.10 This table presents PRKN-mitophagy Boolean model parameters within three subgroups SWEDD, prodromal, parkinsonism with common language (Cl) effect sizes and targets

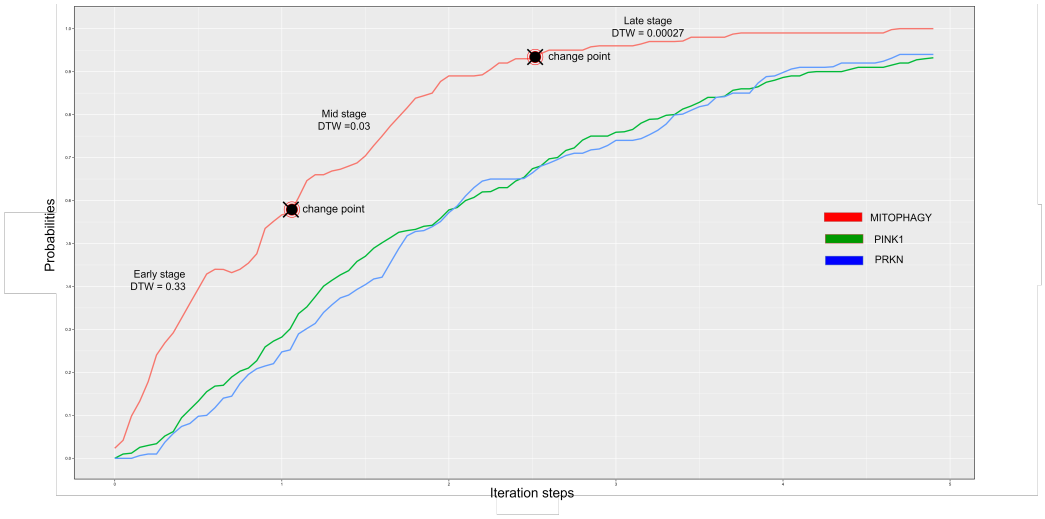


Figure 4.4 The figure shows a representative run of the simulation in PRKN mitophagy model, consisting of 100 iteration steps with 100 repetitions. Dynamic time warping measure similarity between two sequences which differs in speed based on different stages of the simulation (early, mid, and late).

interpreted as early, mid and late stage. The DTW algorithm is then used to calculate the similarities between different time series, which represent the progression of a particular process over time [205]. By using the Pearson correlation coefficient, the pairs of DTW scores were compared to identify which conditions are most correlated across disease subgroups.

Such quantified simulation graphs are generated and summarised for all three PD subgroups, across all considered BMs, their targets and defined outputs. Figure 4.5 summarises the entire workflow.

4.5.2 Model parameterisation using T2DM transcriptomic profile

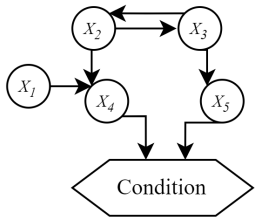
Selected models were further parameterised to study the comorbidity of PD with T2DM. The Differentially Expressed Genes from T2DM datasets (above) were used to introduce perturbations in the models.

For biomolecules that were found to be expressed at lower levels in the transcriptomic dataset, the simulation involved activating or ”overexpression” the biomolecules in silico. In contrast, for genes that were found to be downregulated, the simulation involved in silico knocking out the biomolecule.

The following example in Table 4.11 shows parameterisation of PARKIN pathway, including two sets of parameters (miRNA based and T2DM specific

Inputs: Cohort specific Probabilistic Boolean models

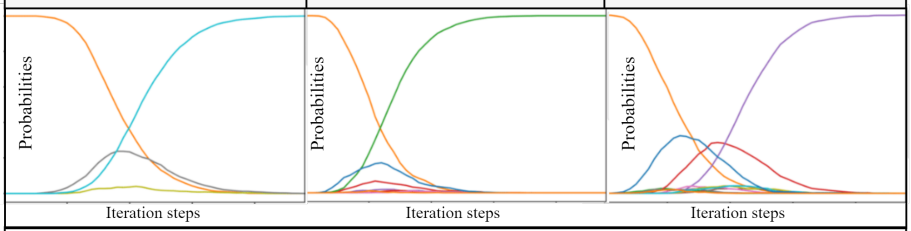
Prodromal	SWEDD	Parkinsonism
$X_3 = P_1$ $X_4 = P_2$	$X_2 = P_3$ $X_5 = P_4$	$X_1 = P_5$ $X_2 = P_6$



Outputs :

A) Conditions probabilities in disease cohorts

Prodromal	SWEDD	Parkinsonism



B) DTW score of a condition probabilities in simulation stages

Stages	Prodromal	SWEDD	Parkinsonism
Early	0.2	0.6	0.3
Mid	0.1	0.3	0.2
Late	0.06	0.01	0.02

Figure 4.5 The figure illustrates an example of a Probabilistic Boolean simulation run using pyMaBoSS tool. The inputs are cohort-specific probabilistic models of prodromal-SWEDD-parkinsonism with X nodes that have different initial states probabilities (P). The outputs represented on two panels: Panel A displays the output of the model in a simulation graph, which presents the probabilities of the outputs in iteration steps. Panel B shows DTW alignment scores, which measure the similarity between the probabilities of the different conditions across the early, mid, and late simulation stages.

which are combined) during the simulation. The miRNA-based parameters were used to represent the cohort specific paramenters. On the other hand, T2DM-specific parameters were used to represent the effects of T2DM on PD cohort progression.

Cohort	Cohorts -based targets	CL-effectsize	T2DM-specific targets (Knockouts)
SWEDD	ATXN3	0.966761	SNCA
	BAG4	0.712208	BCL2
	FBXW7	0.852718	BNIP3
	GABARAPL1	0.92637	
	TIMM17A	0.981296	
	ULK1	0.712208	
Prodromal	ATXN3	0.956198	
	BAG4	0.951438	
	FBXW7	0.960511	
	GABARAPL1	0.960511	
	TIMM17A	0.574494	
	ULK1	0.951438	
Parkinsonism	ATXN3	0.976557	
	BAG4	0.896051	
	FBXW7	0.914474	
	GABARAPL1	0.938137	
	TOMM70	0.942465	
	ULK1	0.896051	
	VPS13C	0.964472	

Table 4.11 The table shows an example of parameterisation for PARKIN pathway, including two sets of parameters (miRNA based and T2DM specific which are combined) during the simulation

4.6 Cohort specific simulation results

Following the parameterisation as discussed above, the selected BMs were analysed in two aspects:

- Molecular mechanisms in specific cohorts: parkinsonism, SWEDD and prodromal
- Role of T2DM comorbidity in these cohorts

For each of the selected models (Table 4.9) two sets of results were obtained: cohort-specific and comorbidity-specific results.

4.6.1 Dopamine transcription

Simulation results of the ”Dopamine transcription” BM show that during the early stages, the SWEDD and prodromal-related groups have similar levels of activation of mitochondrial biogenesis and dopamine metabolism. However, the parkinsonism group has higher levels of activation for these processes. All of the groups also show differences in neuron survival, with lower activity in the SWEDD and prodromal groups (Table 4.12).

As the simulation progresses to the mid and late stages, an increase in activation levels for these processes is observed in all groups except the diabetic group,

		Cohort miRNA			T2DM transcriptomics		
Stage	Conditions	Prodromal	SWEDD	Parkinsonism	Prodromal	SWEDD	Parkinsonism
Early	Mitochondrial biogenesis	0.6771	0.6512	0.3853	0.3170	0.3185	0.1937
	Dopamine metabolism	0.4592	0.4565	0.3856	0.1359	0.1322	0.1238
	Neuron survival	0.3743	0.4894	0.3818	0.3413	0.3274	0.0668
Mid	Mitochondrial biogenesis	0.1693	0.1623	0.0493	0.8890	0.8927	0.0456
	Dopamine metabolism	0.0973	0.0680	0.0496	0.0935	0.2073	0.0436
	Neuron survival	0.0761	0.0906	0.0674	0.9226	0.8951	0.0268
Late	Mitochondrial biogenesis	0.0311	0.0482	0.0002	0.9676	0.9698	0.0023
	Dopamine metabolism	0.0132	0.0135	0.0098	0.0364	0.0765	0.0198
	Neuron survival	0.0110	0.0210	0.0295	0.0364	0.0765	0.0198

Table 4.12 The table presents the DTW scores in dopamine transcription BM for three different simulation stages of prodromal, SWEDD and parkinsonism. The scores are based on three different disease conditions: mitochondrial biogenesis, dopamine metabolism, and neuron survival.

where the levels of activation for mitochondria and neuron survival remain unchanged. During the mid and late stages, one can observe significant difference in mitochondrial biogenesis between the SWEDD/prodromal related T2DM groups and the other groups (Table 4.12). In the parkinsonism-T2DM group, the levels of activation for ”mitochondrial biogenesis” and ”neuron survival” were found to be lower than in the other groups at all stages of the simulation. These findings suggest that there may be a link between diabetes and reduced levels of mitochondrial biogenesis and neuron survival, particularly in cases of parkinsonism-T2DM.

4.6.2 Wnt-PI3K/AKT signalling

The simulation shows that in the early stages, there is a significant difference in insulin resistance between the prodromal group and the SWEDD and parkinsonism groups. The SWEDD and parkinsonism groups have similar levels of insulin resistance, while the prodromal cohort has lower levels of insulin resistance. Both the SWEDD and prodromal groups also have a similar probability of insulin resistance due to T2DM. The active forms of TFEB (phosphorylated TFEB and TFEB SNCA) have a low probability, while the inactive form (TFEB complex) has a high probability of development(Table 4.13).

		Cohort miRNA			T2DM transcriptomics		
Stage	Conditions	Prodromal	SWEDD	Parkinsonism	Prodromal	SWEDD	Parkinsonism
Early	TFEB phosphorylated	0.9503	0.9215	0.9216	0.9032	0.9053	0.8998
	Insulin resistence	0.6341	0.3364	0.3310	0.1111	0.1118	0.0190
	TFEB SNCA complex	0.9622	0.9422	0.9422	0.8964	0.9215	0.9088
	TFEB complex	0.3624	0.3572	0.3580	0.05755	0.0575	0.0125
	Neuron death	0.3090	0.2714	0.2725	0	0	0.0366
Mid	TFEB phosphorylated	0.9503	0.9215	0.9216	0.8847	0.8906	0.8758
	Insulin resistence	0.6341	0.3364	0.3310	0.0213	0.0203	0.0098
	TFEB SNCA complex	0.9622	0.9422	0.9422	0.9080	0.9065	0.9078
	TFEB complex	0.3624	0.3572	0.3580	0.0007	0.0006	0.0006
	Neuron death	0.3090	0.2714	0.2725	0	0	0.0048
Late	TFEB phosphorylated	0.9298	0.9276	0.9243	0.8997	0.9009	0.8991
	Insulin resistence	0.0241	0.0222	0.0212	0.0213	0.0203	0.0098
	TFEB SNCA complex	0.9324	0.9279	0.9208	0.9066	0.8914	0.8931
	TFEB complex	0.0296	0.0098	0.0097	0	0	0
	Neuron death	0.0192	0.0005	0.0009	0	0	0

Table 4.13 The table presents the DTW scores in Wnt-PI3K/AKT BM for three different simulation stages of prodromal, SWEDD and parkinsonism. The scores are based on different disease conditions: TFEB phosphorylated, Insulin resistence, TFEB complex and Neuron death.

4.6.3 FOXO3 activity

The simulation shows that in the early stages, the prodromal-related cohort develops ”Fission Fusion” endpoint at a higher rate than the other groups, and this trend continues in the mid and late stages. The prodromal-T2DM cohort also has faster development of autophagy and responses to oxidative stress compared to the other groups. The prodromal and SWEDD groups have similar responses to oxidative stress, which are greater than those in the parkinsonism group. In terms of autophagy, the SWEDD group has a lower rate than the prodromal and parkinsonism groups, which are similar to each other. The parkinsonism cohort has a higher probability of apoptosis compared to the other groups (Table 4.14).

At the mid stage of the simulation, the probability of fission and fusion in SWEDD and parkinsonism is similar. Further, autophagy develops significantly more in patients with parkinsonism than in other cohorts. At the mid stage, the probability of apoptosis becomes similar in prodromal and SWEDD. In SWEDD, the probability of autophagy and apoptosis increases in the late stage compared to the prodromal and parkinsonism cohorts (Table 4.14).

At the mid- and late stages, the SWEDD group has a significantly higher probability of oxidative stress compared to the prodromal and parkinsonism groups. The SWEDD-T2DM group also exhibits an increased response to oxidative stress and an increased probability of mitochondrial biogenesis. The prodromal cohort also has a high probability of mitochondrial biogenesis. Additionally, the

parkinsonism group's mitochondrial biogenesis is less affected by T2DM than in the other groups (Table 4.14).

		Cohort miRNA			T2DM transcriptomics		
Stage	Conditions	Prodromal	SWEDD	Parkinsonism	Prodromal	SWEDD	Parkinsonism
Early	Response to oxidative stress	0.6331	0.6220	0.6949	0.4473	0.4850	0.5954
	Fission Fusion	0.6361	0.8529	0.8111	0.4676	0.6192	0.6858
	Autophagy	0.5561	0.6491	0.5514	0.3789	0.3962	0.4145
	Mitochondrial biogenesis	0.5622	0.5419	0.5824	0.2045	0.1843	0.1500
	Apoptosis	0.7547	0.7707	0.5508	0.3947	0.4657	0.3255
Mid	Response to oxidative stress	0.5138	0.4952	0.5708	0.5133	0.4938	0.6417
	Fission Fusion	0.5413	0.6091	0.6015	0.5188	0.7658	0.6431
	Autophagy	0.5096	0.5418	0.4814	0.5195	0.4671	0.5028
	Mitochondrial biogenesis	0.1257	0.1465	0.1633	0.0744	0.1251	0.0777
	Apoptosis	0.4666	0.4640	0.3315	0.4231	0.4450	0.4302
Late	Response to oxidative stress	0.5099	0.4577	0.5485	0.5325	0.4703	0.6460
	Fission Fusion	0.5080	0.5133	0.5469	0.5528	0.6010	0.6156
	Autophagy	0.5399	0.4978	0.5367	0.5323	0.4878	0.5567
	Mitochondrial biogenesis	0.0208	0.0372	0.0712	0.0197	0.0436	0.0392
	Apoptosis	0.3177	0.2623	0.2752	0.2664	0.2940	0.3221

Table 4.14 The table presents the DTW scores in FOXO3 BM for three different simulation stages of prodromal, SWEDD and parkinsonism. The scores are based on five different disease conditions: Response to oxidative stress, Fission Fusion, Autophagy, Mitochondrial biogenesis and Apoptosis.

4.6.4 mTOR-MAPK signalling

The simulation shows that in the early stages, there are significant differences in the activity levels of glycolysis and catabolism among all cohorts. The change in glycolysis begins earlier in the SWEDD group and continues to have higher activity in the late stage. The SWEDD-T2DM group has lower levels of glycolysis at all stages of T2DM, while catabolism increases whenever glycolysis decreases (Table 4.15). In the parkinsonism group, catabolism is more active at the early and mid stages compared to the SWEDD and prodromal groups. Despite the increased activity of glycolysis, the SWEDD group has a higher probability of catabolism in the mid/late stages (Table 4.15). As the simulation of catabolism continues into the late stages, a significant difference is observed between the parkinsonism/SWEDD cohort and the prodromal cohort. Additionally, the parkinsonism-T2DM group has higher levels of glycolysis activity than the other groups with T2DM (Table 4.15).

		Cohort miRNA			T2DM transcriptomics		
Stage	Conditions	Prodromal	SWEDD	Parkinsonism	Prodromal	SWEDD	Parkinsonism
Early	Glycolysis	0.8151	0.7664	0.9058	0.3040	0.3971	0.1740
	RHEB lysosome	0.6653	0.3861	0.6113	0.1382	0.0098	0.0286
	AKT	0.4171	0.3893	0.4172	0.0630	0.0122	0.0081
	Catabolism	0.8087	0.8350	0.7342	0.3553	0.4528	0.1848
Mid	Glycolysis	0.3077	0.3461	0.5149	0.3365	0.3690	0.2926
	RHEB lysosome	0.2368	0.0937	0.2166	0.0744	0.0098	0.0048
	AKT	0.1115	0.0962	0.1930	0.0555	0.0098	0.0065
	Catabolism	0.3688	0.3193	0.2579	0.3458	0.2822	0.3093
Late	Glycolysis	0.1493	0.1397	0.1780	0.1666	0.1799	0.1458
	RHEB lysosome	0.1227	0.0446	0.0643	0.0492	0.0002	0.00003
	AKT	0.0427	0.0296	0.0439	0.0128	0.0001	0.00001
	Catabolism	0.1601	0.1392	0.1398	0.1924	0.1629	0.2083

Table 4.15 The table presents the DTW scores in mTOR BM for three different simulation stages of prodromal, SWEDD and parkinsonism. The scores are based on four different disease conditions: Glycolysis, RHEB lysosome, AKT and Autophagy.

4.6.5 PRKN mitophagy

At all stages, there is a significant difference in the activation of mitophagy across cohorts. Mitophagy activation begins more quickly in the early stages of the SWEDD and prodromal groups compared to the parkinsonism group. The parkinsonism group has a higher probability of PINK1 accumulation than the other groups. T2DM decreases the probability of mitophagy in the parkinsonism group (Table 4.16).

		Cohort miRNA			T2DM transcriptomics		
Stage	Conditions	Prodromal	SWEDD	Parkinsonism	Prodromal	SWEDD	Parkinsonism
Early	Mitophagy	0.3303	0.3303	0.4719	0.0537	0.0537	0.0422
	PRKN ubiquitinated	0.6165	0.6165	0.4545	0.2041	0.2041	0.0643
	PINK1 accumulation	0.5952	0.5952	0.5870	0	0	0.1556
Mid	Mitophagy	0.0325	0.0325	0.0820	0.0078	0.0078	0.0335
	PRKN ubiquitinated	0.1918	0.1918	0.1009	0.1065	0.1065	0.1212
	PINK1 accumulation	0.1982	0.1982	0.1314	0	0	0.0658
Late	Mitophagy	0.0002	0.0002	0.0129	0	0	0.0101
	PRKN ubiquitinated	0.0644	0.0644	0.0933	0.0327	0.0327	0.0534
	PINK1 accumulation	0.0747	0.0747	0.0196	0	0	0.0204

Table 4.16 The table presents the DTW scores in PRKN mitophagy BM for three different simulation stages of prodromal, SWEDD and parkinsonism. The scores are based on three different disease conditions: Mitophagy, PRKN ubiquitinated, and PINK1 accumulation.

4.6.6 Similar characteristics in the disease subgroups

Understanding the similar characteristics of the disease conditions in the disease subgroups is an important aspect of diagnosis and treatment planning, as it can help to inform the selection of appropriate therapies and interventions. The DTW algorithm identified and synthesised the similarities between disease conditions in subgroups (see Section 3.4.3). A lower DTW score between two conditions in different subgroups indicates a higher degree of similarity in the progression of these conditions over time, while a higher DTW score indicates more differences in the progression. By using the Pearson correlation coefficient (see Section 3.4.3, the pairs of DTW scores were compared to identify which conditions are most correlated, and to understand the patterns of their progression over time. The Pearson correlation coefficient identified the highly correlated DTW similarity score. The following findings are selected based on a correlation coefficient greater than 0.98, indicating a strong positive correlation (Table 4.17). The results show that in the early stages of simulations, most of the observed conditions are correlated to dysregulation of mitochondria. These conditions continue to be present in the mid and late stages, manifesting as apoptosis, catabolic processes, dopamine metabolism, and neuron death (Table 4.17). At the early stage of the simulation, all of the cohorts are affected by T2DM. T2DM has a similar effect on dopamine metabolism and neuron death in the late stage of simulation (Table 4.17). The prodromal and SWEDD groups have similar levels of mitochondrial biogenesis at the early and late simulation stages of T2DM. There are no differences in mitophagy activity between the simulation stages, but insulin resistance activity is similar only in the early stages of the simulation (Table 4.17). The SWEDD and parkinsonism groups have similar probabilities of insulin resistance activation in the early stage of simulation, fission metabolism in the mid stages, and catabolism in the late stages. In the early and late stages of simulation, the prodromal and SWEDD groups have similar probabilities of mitophagy activation, while dopamine metabolism appears similar in the late stage. The neuron survival condition is similar in the prodromal and parkinsonism groups in the early stage of the simulation, and the autophagy pathway is similar in the early and late stage (Table 4.17).

	Conditions in stage		
Disease subgroups	Early	Mid	Late
Prodromal+T2DM, SWEDD+T2DM	Insuline resistence Mitophagy Mitochondrial biogenesis	Insuline resistence Mitophagy Apoptosis	Insuline resistence Mitophagy Mitochondrial biogenesis
SWEDD+Parkinsonism	Insuline resistence	Fission fusion	Catabolism
Prodromal+SWEDD	Mitophagy	–	Mitophagy Dopamine metabolism
Prodromal+Parkinsonism	Autophagy Neuron survival	–	Autophagy
T2DM (all cohorts)	Dopamine metabolism	–	Neuron death

Table 4.17 The table presents the highly correlated disease conditions within disease subgroups (ranging from 98%-100%) at various stages of simulation using the Pearson correlation coefficient.

Chapter 5

Discussion

Neurodegenerative diseases, including Parkinson's disease (PD), are conditions that arise from the complex interaction of genetic and environmental factors and are characterized by disruptions in multiple molecular and cellular processes [309]. These conditions are associated with a wide range of symptoms and can be further complicated by comorbidities. The need to understand and effectively treat these conditions is growing as their prevalence increases [310]. To better understand the underlying molecular and cellular mechanisms of these diseases and identify more precise and effective therapeutics, new approaches and technologies are needed, including systems biology [311]. Systems biology is a holistic approach that focuses on the interactions between various biomolecules of a biological system and how they contribute to the function and behavior of the system as a whole [312]. Systems biology approaches often involve the use of computational models and the integration of large datasets to gain a comprehensive understanding of the system [313, 314]. The goal of these approaches is to improve our understanding of the underlying causes of complex diseases and develop more precise and effective therapies.

Modelling biological processes is a crucial tool for understanding the complex interactions and relationships within biological systems. There are different modelling approaches that can be used, such as static network analysis and dynamic modelling. Static network analysis can be useful for identifying correlations and clusters between biomolecules [76], while dynamic modelling focuses on simulating the causal relationships between biomolecules and their measurement values [315, 313]. Dynamic modelling approaches include the use of Petri nets, ODEs, Boolean, and multi-valued models. Logical models such as Boolean or multi-valued discrete models are interesting alternatives as they do not require detailed kinetic information. However, these types of models may not be suitable for certain research areas, such as pharmacogenomics, where insufficient data is a limitation. The development of Boolean models requires a combination of

large-scale studies, data analysis, and a thorough understanding of the underlying biology of the disease [77].

The study used Boolean modelling to investigate the complexity of PD by simulating the dynamic interactions between various biomolecules. The BMs are used to test hypotheses about the role of specific biomolecules in the development of PD by simulating the effect of these biomolecules on the disease progression and evaluating the results against experimental or observational data. Additionally, these models are used to explore the impact of multiple perturbations on the disease and identify patterns that may not be apparent from experimental or observational data alone. Furthermore, the use of stratified BMs can aid in the diagnosis of PD by analyzing the specific factors that contribute to the disease progression and tailoring treatment strategies to the specific needs of each subgroup of patients. Overall, BMs can provide a valuable computational tool that can aid researchers in understanding the complexity of PD, developing new diagnostic criteria, targeted interventions, and better treatment strategies, ultimately improving care for patients with PD.

5.1 Constructing Boolean models from knowledge repositories

The BMs were constructed in an automatic fashion using CaSQ tool[93]. One of the key advantages of using the CaSQ tool to construct BMs from PD map diagrams is its ability to apply specific rewriting rules to simplify the model and make it more manageable. This is demonstrated in the method section (see Chapter 3), where the Process Description notation is reduced to the Activity Flow notation, and logical functions and interactions are inferred from the Process Description (see Figure 3.1). This reduction process is important for making the model more tractable for downstream analysis and modeling.

Another advantage of using the CaSQ tool is its ability to translate diagrams into the SBML-qual format, which is a widely adopted standard for representing qualitative dynamic systems in the systems biology community. SBML-qual allows the description of the structure of a model (the biomolecules and interactions) as well as the mathematical equations that describe how these components behave over time. The use of this format allows for easy sharing and comparison of models with other researchers and provides a standardized way of representing models.

The transformation of the diagrams to Simple Interaction Format (SIF) using CaSQ tool added advantage, as it allowed for the use of a range of tools to analyse the diagrams. This is important in the context of understanding complex

diseases such as Parkinson's disease, where a comprehensive understanding of the complex mechanisms is necessary. By using different tools to analyse the diagrams, researchers can gain a more holistic view of the disease, which can improve the understanding of the underlying mechanisms.

The reduction of Process Description notation into Activity Flow notation and inferring logical functions from interactions, combined with the use of SBML-qual format, allowed for the creation of models that are both accurate and computationally efficient. This makes the generated models more suitable for studying large and complex systems as well as allows for a deeper understanding of the system and more accurate predictions about its behavior. Therefore, the constructed BMs have the potential to be scaled up in order to study larger and more complex systems.

It is worth noting that the use of specific knowledge disease resources, such as PD map, has the advantage of providing a high quality of knowledge specificity. This is because the diagrams are focused on a specific disease and are created and reviewed by experts in the field. This high level of expertise and specificity makes the generated models more reliable.

5.2 Data-driven model parameterisation and simulations

5.2.1 Comparison of simulation algorithms

This study highlights the importance of considering different algorithms for different goals and characteristics of the systems being studied. The Heuristic and SAT solver algorithms were found to be the most promising options for downstream analysis as they achieve a good trade-off between performance and biological relevance. The results can be valuable for researchers who are interested in using simulation algorithms to analyse biological systems. The performance of different simulation algorithms were compared for understanding the characteristics and calculating the performance of BMs. By comparing asynchronous and synchronous simulations in selected models, the aim was to understand how their characteristics are affected by different updating schemes. The attractor analysis showed that the state trajectories converge to fixed or cyclic attractors under different updating schemes. In this context, four algorithms (HyTarjan, Heuristic, Decomp, and SAT) in terms of their reachability in pathways. The data indicated that in the case of asynchronous solutions, the HyTarjan algorithm outperformed the Heuristic approach. When comparing Decomp vs SAT, the data suggest that there was not a single solution that consistently outperformed the other. The attractors produced by these methods in terms of their biological rele-

vance showed that both the HyTarjan and Decomp solutions produced attractors that were not biologically relevant for all the considered pathways. In contrast, the Heuristic and SAT solver algorithms produced attractors that were viable for all the pathways except the Wnt/PI3K pathway. The reason for this is that the HyTarjan and Decomp algorithms use an aggressive decomposition which overly fragments the resulting models. Therefore, the Heuristic and SAT solver algorithms were selected for downstream analysis.

5.2.2 Parameterisation using omics data

The study explored the use of BMs and probabilistic BMs as a modelling framework for understanding the dynamics of biological systems. The probabilistic BMs can be adjusted to incorporate specified levels of model parameters, allowing modelling systems with varying probabilities based on expression levels. This allowed to better capture the inherent variability in biological systems. Probabilistic BMs are stochastic in nature, meaning that the outcome of the simulation is not always deterministic. This may be relevant more for the biological systems that exhibit stochastic behaviour such as expression levels of the biomolecules. Such a model could be used to simulate the effects of different medications or interventions on the activity of biomolecules involved in a disease, and predict how these interventions may affect the disease progression in different patient subgroups [316].

A high-quality data set of miRNAs sequenced from the whole blood of PD patients from different subtypes of the disease was obtained from the PPMI study [69]. The existence of microRNAs in the circulation is considered a promising biomarker for both diagnosis and prognosis of disease because of their high stability in human fluids [317, 297]. Differentially expressed miRNAs were evaluated to include only validated interactions between miRNAs and genes in brain tissue. Only stable correlations between miRNA expression and target gene expression were chosen. In addition, the filtered miRNAs were enriched in the PD map as PD-specific knowledge resource and overlapped with substatia nigra targets. As a final step, the expressions of miRNAs are compared with the curated data resources and other experiments related to PD in the literature (Table 4.7). Most of such filtered, significant miRNAs showed downregulation in PD and were involved in mitochondrial dysfunction. The enrichment analysis of the PD map revealed significant pathways that may be involved in disease progression. Using these pathways, probabilistic BMs were created and the initial states are chosen based on the targets of miRNAs biomarkers from PPMI dataset. The probabilities of the initial states are calibrated based on the calculated effect sizes. Boolean simulations were then performed based on the disease subtypes.

Type 2 diabetes mellitus (T2DM), is a significant PD comorbidity, affecting the

progression of different disease subtypes [63]. In this work, the transcriptomic datasets related to T2DM were analysed to understand the relevance of diffrentially expressed genes to PD and T2DM. The data is filtered based on common miRNA-target pairs from the literature and substantia nigra dataset on the PD map(Table 4.8). In this context, upregulated genes were treated as overexpression perturbation and donwregulation as knockout. Such simulation of T2DM as point perturbations separately after simulating a cohort of PD patients, allowed to study the effects of T2DM on PD progression in isolation. This allowed to identify specific interactions and pathways that may be involved in the co-occurrence of these two disorders, and can provide a clearer understanding of the underlying biology.

5.3 Validation of the constructed models

Models were built using systems biology standards, making them interoperable with different tools and programs. Systems biology standard formats (e.g. SBML packages) improved model reproducibility and made pipeline development easier. By using these standard formats, researchers can ensure that their models are easily readable and understandable by other scientists, and that they can be integrated with other models and data in a consistent and reliable way. This improves the reproducibility of models and makes it easier to develop pipelines and workflows that involve multiple models or data sources.

In general, computational models constructed without sufficient detail may result in inaccurate predictions [318]. Thus, a focus on model quality during the construction step is crucial to minimize false positive results during simulation [251, 319]. In this context, the verification step is a critical part of the modelling process, as it helps to ensure that the model is a reliable representation of the system being studied and that it can be used to make accurate predictions and inform decision-making.

In this work, the BMs were evaluated by showing their ability to reproduce experimentally validated studies on different scales of complexity (Section 3.3). Results shown in the Table 4.6 indicate that the simulated behaviour of the models considered for this work is in line with the expected behaviour based on different scenarios, reproducing expected behaviour for both original and perturbed conditions. Simulation of known perturbations proves the models' capability to recreate known pathological conditions, helps to understand them, and allows to prioritise a reliable set of biomolecules and translation rules.

The selection of PD map diagrams for downstream modelling and verification is an important step to explain the underlying biological mechanisms of PD and provide targets for the development of therapies. These diagrams were

chosen based on their relevance to PD phenotypes, such as mitochondrial dysfunction, dopamine dysregulation, alpha-synuclein aggregation, neuroinflammation, oxidative stress [320, 321].

5.4 Structural and functional validation of the Boolean modelling approach

5.4.1 Modelling of the TCA cycle

The model of the TCA cycle pathway from the PD map was used to validate the modelling approach. The TCA cycle is a well-studied pathway with a wealth of experimental data available, including measurements of enzyme activity and metabolite levels. This allowed to compare the BM predictions with experimental data to determine the accuracy and reliability of the model. Moreover, desregulation of the TCA cycle is linked to oxidative stress, inflammation, and cell death, which are all hallmarks of PD [208, 212, 209].

To investigate the TCA cycle BM, the structural and functional roles of key molecules were analysed. This included examining the involved enzymes and cofactors, as well as the reactions they catalyze and the intermediates they produce. The regulatory mechanisms that control the activity of these molecules were considered to predict impact on the TCA cycle [322]. The effects of overexpression and knockout of regulatory mechanisms that control the activity of the enzymes in the TCA cycle were modelled in the TCA cycle BMs. In particular, the results from the structural analysis were used to calculate the sensitivity of the TCA cycle to knockouts and overexpressions. This involved simulating knockouts and overexpressions of a particular biomolecule on the overall activity of the TCA cycle. To verify the results, literature was reviewed to find experimental studies that perturbed TCA cycle molecules in model organisms, such as yeast or mice [323, 324].

Selected perturbations in the Boolean model resulted in the following findings supported by the literature (Table 4.6):

- Activation of acetyl CoA, NADH, and PDKs in silico increased the phosphorylation reaction, which reduced the activity of the pyruvate dehydrogenase complex (PDC) [208]. This led to a decrease in succinic semialdehyde and succinic acid [209].

- Simulated KGDHC knockout predicted deficiency in succinic acid and succinyl CoA and a downstream decrease in ATP and GTP [210, 211].

- Simulated isocitrate dehydrogenase knockout lead to a decrease in ATP production and inhibited the oxidative decarboxylation of isocitrate [212]. Additionally, levels of oxoglutarate, an intermediate component that serves as the carbon skeleton for nitrogen assimilation, are observed to decrease in response to L-glutamate [212, 325].

- Knockout of SIRT3 downregulated oxoglutaric acid, glutamate hydrogenase 1 (GLUD), and ATP levels, which has a direct impact on mitochondrial function [206]

- Finally, simulation of the effect of alpha KGDH and GTP, GDP on ATP levels was validated using probabilistic BM based on temporal-fluxomics data that describe the oscillations of metabolites in the TCA [207]. The simulated activity levels matched the measured concentrations in synchronized HeLa cells at various time points post-release from growth arrest (Table 4.4).

While the TCA model used in this study does not answer the same research questions as quantitative models, it reproduces molecular activity that could regulate ATP levels and mitochondrial function. These findings suggest that the TCA BM is a reliable tool for describing the general overview of the energy metabolism and may provide insight into the mechanisms underlying the oscillations of cellular metabolites. This evaluation process has certain limitations. For example, the literature search may not capture all relevant data, and the simulations may not accurately reflect all the complex interactions that occur in cells. It is also important to consider the specific context in which the TCA cycle is being evaluated, as it may vary depending on the organism or tissue being studied. However, the main focus is validating BMs of the TCA cycle to ensure that the model accurately reflects the underlying biological processes.

5.4.2 Modelling of the signalling pathways

Dopamine transcription: The role of the NR4A2 gene

The NR4A2 gene is a key player in the regulation of dopamine levels and the development and maintenance of neurons. This gene is expressed in a variety of tissues, including the brain, and is shown to be involved in several important biological processes, such as neuron survival, mitochondrial biogenesis, and apoptosis [326]. The in silico knockout of NR4A2 correctly predicts inhibition of BDNF production [213] and regulation of expression of several other genes involved in neurotransmitter metabolism and transport, including dopamine [216, 217, 218]. Activation of the SIRT1 gene, another gene involved in neurotransmitter metabolism, improves mitochondrial biogenesis, a process that is necessary for the production of new mitochondria.

Wnt-PI3K/AKT signalling: Implications for dopaminergic neurogenesis and autophagy

The Pi3K/AKT pathway and Wnt signaling are critical for dopaminergic neurogenesis and are involved in important developmental processes and in the aging process, which is a major risk factor for PD. These two signaling pathways also have common downstream targets, suggesting the potential for crosstalk and synergistic therapeutic approaches [44, 327]. Both pathways can regulate the activity of certain signaling proteins, such as the protein GSK3B [328, 229]. Moreover, they regulate pathways of autophagy, protein translation, and their role in neuronal survival. The following findings validated the contents and the structure of the patway.

- Effect of LRRK2: LRRK2 gene its G2019S mutant form is associated with an increased risk of developing PD [219]. Overexpression of the LRRK2 G2019S mutant enhanced autophagy, a process that involves the degradation and recycling of cellular biomolecules. This effect appears to be mediated by the inhibition of mTORC1/2, reflecting the reported findings [220, 125, 222, 223]. Modelling overexpresion of protein AMPK predicted reactivation of mTORC1 and is consistent with proposed interaction between amino acid sensing, mTORC1 signaling, and autophagy [221].

- Regulation of autophagy and protein translation: TFEB is a transcription factor that plays a crucial role in the regulation of the autophagy [225]. Simulated TFEB overexpression increased the activity of autophagy by 36.9%, leading to a neuroprotective effect. However, TFEB also activates protein synthesis inducers such as EIF4E and RPS6KB1 (Table 4.5). These complex effects of TFEB on the autophagy are still under investigation, however modelling results in this regard reflect recent findings [224, 237, 225]. Interestingly, joint activation of TFEB and protein synthesis inhibitor 4EBP2 resulted in a 10-13% increase in autophagy activity compared to their single activations. Further, the inhibition of RPS6KB1 and PHLPP and activation of TFEB significantly decreased neuronal death (by 96.3%) and the activated autophagy (Table 4.5).

- Effects of Wnt and GSK3B: Simulated overexpression of Wnt protein and the inactivation of GSK3B increased autophagy, confirming reported findings [329, 330, 331]. In turn, simulated GSK3B inhibition and overexpression regulated neurogenesis as reported in [328, 332]. The combination of modulating Wnt signaling and GSK3B indicated the potential towards a useful neuroprotective treatment in the early stages of disease progression (Table 4.5).

Overall, besides confirming the litearture findings, validation study in Wnt/PI3K pathway brought up interesting hypotheses about combined modulation of its elements. However, the relationship between the discussed biomolecules and neurogenesis and autophagy is complex and may depend on specific cell type, activity of other signaling pathways, or the stage of disease progression.

FOXO3 activity: Impact on mitochondrial autophagy

The FOXO3 activity pathway plays a significant role in mitochondrial homeostasis [232]. Simulated FOXO3 activation increased autophagy specifically in mitochondria and led to activation of a number of biomolecules, including BECN1, GABARAPL1, MAP1LC3A, BNIP3, ATG12, and MUL1. These biomolecules are known to be important regulators of autophagy [233], and their activation in response to FOXO3 activation suggests that FOXO3 and these RNA-mediated pathways are closely interconnected in the regulation of autophagy.

5.5 Modelling-based patient stratification by disease subgroup analysis

Having validated the Boolean modelling approach and models based on the PD map, the selected BMs were parameterised with cohort-based data and simulated to stratify the dynamics different molecular mechanisms during disease progression. Simulations of these models provided time series, used to represent stages in the disease progression. This allowed to identify the similarities of the molecular activity across the PD subtypes. The stratification is an important aspect of this work, because its potential to inform treatment decisions and improve our understanding of the underlying mechanisms of the disease.

5.5.1 Specific characteristics in each cohort

The prodromal cohort exhibits molecular changes that may result in higher probabilities of motor signs compared to other cohorts, possibly related to impaired neuronal autophagy [333]. This may also explain the higher levels of autophagy and responses to oxidative stress seen in prodromal cohorts. The models show that in the prodromal cohort mitochondrial turnover is more frequent (”Fission and Fusion” output) with higher probabilities than other cohorts in the early stages of simulation, and this pattern continues with a higher probability in the mid and late stages of simulation (Table 4.14). This finding is consistent with previous research suggesting that mitochondrial abnormalities may occur early in the course of PD [47, 334]

In the SWEDD cohort, an interesting aspect is observed inhibition of "Glycolysis and catabolism" output in the mTOR-MAPK signalling model. At the same time the protein RHEB, which has a neuroprotective role, is highly active. This suggests that RHEB may play a role in decreasing catabolic processes and potentially protecting against the development of PD [335](Table 4.15). In the later stages of SWEDD, there is an increase in catabolism, which is the breakdown of molecules to release energy. This increase in catabolism is accompanied by increased glycolysis activity, which is the breakdown of glucose to produce energy.

In parkinsonism, dopamine transcription and Wnt-PI3K/AKT models show that mitochondrial biogenesis and dopamine transcription change rapidly with lower change points in the mid and late stages simulation. This finding may be due to the fact that the parkinsonism syndrome tends to progress more rapidly than other PD subgroups [336].

5.5.2 Characteristics of prodromal and SWEDD subjects

The SWEDD and prodromal cohorts are similar in dopamine metabolism or mitochondrial biogenesis in the early stages of PD. The probabilities of both phenotypes is nearly identical, as demonstrated in (Table 4.12). This finding suggests that the early stage of prodromal and SWEDD may refer to a stage at which individuals do not fulfill diagnostic criteria for clinical PD. However, a recent study propose that SWEDD patients do not have early PD [337].

Dopamine transcription and Wnt-PI3K/AKT models suggest that neuronal activity may be influenced by dopamine metabolism and mitochondrial biogenesis, as these processes are important for energy production and the function of neurons. Specifically in Wnt-PI3K/AKT, SWEDD and prodromal conditions show lower levels of neuronal activity compared to parkinsonism in early stages. It is possible that dopamine metabolism is sustained in the SWEDD and prodromal early stages of simulation for longer periods of time than in the parkinsonism. As a result, dopaminergic neurons may be affected by oxidative stress, leading to a decrease in their survival rate [45].

Oxidative stress response may be related to dopamine metabolism and neuronal activity because oxidative stress can damage cells and disrupt normal cellular function, including dopamine metabolism and neuronal activity[45]. Prodromal and SWEDD patients exhibit similar oxidative stress responses that are higher than those observed in parkinsonism patients. This may explain the slight differences in dopamine metabolism and lower neuron survival activity in the dopamine transcription pathway observed between these two subtypes and other conditions in early stages of simulation.

Glycolysis and catabolism are central processes that are vital for the produc-

tion of energy in cells. Dysregulation of these processes is observed in a wide range of disease states, including PD. Simulation results in both cohorts show that changes in glycolysis and catabolism occur earlier in SWEDD and prodromal, compared to parkinsonism (Table 4.15).

Mitophagy is the process of degrading and recycling mitochondria, and changes in this process may affect the function and survival of mitochondria and cells. Dysregulation of mitophagy is implicated in the development of the prodromal and SWEDD (Table 4.16). In the PRKN mitophagy model, the prodromal and SWEDD cohorts show higher probability of mitophagy activation than those with parkinsonism. This activation in prodromal and SWEDD is mediated by the protein ULK1, suggesting that the process may be independent of PRKN [338], despite higher PINK1 accumulation in the simulation for parkinsonism cohort.

5.5.3 Characteristics of T2DM-related cohorts

Diabetes-parameterised and cohort-specific models demonstrate a series of differences from the results discussed above. One of the most affected is the Dopamine transcription model. It features significantly lower levels of mitochondrial biogenesis and neuronal survival at mid and late stages of simulations (Table 4.12). This is in line with a recent study, linking T2DM to a decline in neuron survival, mitochondrial biogenesis, and dopamine metabolism, where T2DM was associated with oxidative stress and decreased levels of dopamine and its metabolites in the striatum [339, 340]. Interestingly, in the mid and late stages, the probability of dopamine metabolism activation differs among the cohorts. In particular, that dopamine metabolism is less affected in SWEDD-T2DM cohort (Table 4.12).

For the Dopamine transcription model parameterised for parkinsonism-T2DM cohort, in the mid and late stages of the mitochondrial biogenesis is less impacted compared to other T2DM cohorts (Table 4.12). However, in the early stages of the disease, T2DM is found to increase the cellular response to oxidative stress, potentially through the activation of quality control mechanisms such as Autophagy and Fission and Fusion. These processes may increase apoptosis, a form of cell death triggered in response to cellular stress. Moreover, in mTOR-MAPK signalling model for parkinsonism-T2DM, higher glycolysis activity was observed. Coincidentally, in T2DM there is also an increase in the inactivated form of RHEB and the activation of anaerobic glycolysis. This shift towards anaerobic glycolysis is thought to occur as the brain tries to maintain ion homeostasis by providing a limited amount of energy through the breakdown of glucose in the absence of oxygen. However, this process ultimately leads to chemical changes that result in cell death [341](Table 4.15). Finally, the probability of mitophagy is decreased in parkinsonism-T2DM cohort compared to those without T2DM. The results highlight the increased activation probability of the protein VPS13C

in T2DM, which delays the progression of mitophagy. In support of this, two novel cases are reported of patients who developed dementia and early onset parkinsonism in the absence of VPS13C [342].

5.5.4 Common characteristics in all cohorts

The results show that insulin resistance is a common feature in PD, and its dysregulation is observed in all disease subgroups. Insulin resistance is a condition in which the body's cells do not respond properly to the hormone insulin, leading to high blood sugar levels and an increased risk of diabetes and other health problems [343].

The models suggest that the development of insulin resistance may be linked to the activity of the transcription factor TFEB (Table 4.13). The BMs show that the active forms of TFEB tend to have low activity [344], while the inactive form of TFEB, found in a complex with the 14-3-3 protein in the cytoplasm, tends to be elevated [345]. The 14-3-3 proteins are a family of highly expressed brain proteins with neuroprotective effects in multiple PD models [345]. However, high levels of the inactive form of TFEB may suggest a decrease in 14-3-3 protein, which may increase the aggregation of alpha-synuclein and impair cellular processes, leading to insulin resistance [343]. Recent studies show that the use of antidiabetic drugs has a beneficial role to control PD symptoms [346, 347, 348]. One of these drugs is Metformin, suggested as a neuroprotective drug in the prodromal cohort. Metformin can reduce alpha-synuclein aggregation and improve cellular processes associated with age-related conditions [348]. The model suggests that dysregulation of TFEB and its regulated genes plays an important role in insulin resistance and controlling mitochondrial function in PD. A recent study shows that abnormalities in TFEB cause a failure of endolysosomal and autophagic pathways[349].

5.6 Application of results in translational research

The aim of translational medicine is to bridge the gap between basic research and clinical domain. Modelling and simulation techniques facilitate this goal by using models build based on existing knowledge from bench experiments and disease-relevant omics datasets to develop new hypothesis to understand the diseases, and propose better therapies and diagnostics. In line with this goal, the Boolean modelling approach is rooted in systems-biology and systems-medicine, which allows to iteratively improve our understanding through a continuous cycle of data-driven modelling and model-driven experimentation. The model simulations

can be used to generate testable hypotheses, and once these hypotheses are validated through experiments, one can incorporate the new knowledge into the models, thereby improving their accuracy.

One of the key benefits of this approach is its potential to propose therapy-related hypotheses. Boolean modelling approach can be used to design perturbation experiments by comparing the model attractors to the disease signature. This allows to identify the basins of attraction that could alleviate the pathological signature of the disease and suggest the best combinations of targets to reach a healthy state.

For instance, the study of the Wnt/PI3K pathway yielded valuable insights and proposed hypotheses about the combined modulation of its elements. By verifying the model constructed from this pathway, and through a series of simulations, the existing literature findings were not only confirmed, but also uncovered new and intriguing hypotheses that provide a deeper understanding of the complex interactions within this pathway. Additionally, the model proposed that dysregulation of mitophagy varies among different disease cohorts. This indicates that the mitophagy targets and treatment strategies should be tailored to each disease subgroup. To further explore this finding, experimental models that represent each subgroup could be designed and the targets predicted by the model could be perturbed to observe the pathological signature of mitophagy in each subgroup.

The models can be used to improve similarity-based differential diagnosis in PD by identifying the common cross-talk between different subtypes of the disease. PD has different subtypes that may present with similar phenotypes (endpoints), but have different underlying causes. With the help of the model results, the similarity-based differential diagnosis can compare a patient's endpoints to those of similar conditions, including other subtypes of PD, in order to identify the most likely underlying cause.

In summary, this approach aims to provide a tool for understanding the underlying mechanisms of disease, as well as developing new therapies, by providing a clear understanding of the disease-causing dysregulations in biological systems and design tailored strategies to address it.

5.7 Main limitations of the study

5.7.1 Limitations of model construction and training

In this study, construction and parameterisation of a model are two separate processes. First, a static diagram is translated into a dynamic model of a global picture of a disease mechanism. Still, such general model has to be parameterised

for different disease conditions using omics data. Then, the training process with omics data improves our understanding about disease heterogeneity but requires adaptation to suit other conditions. It may be more effective to begin the process of omics integration along with the construction of the model. This would help to fill the gaps of the incomplete information and revise our knowledge about curated interactions.

In another vein, CaSQ was selected as the translation tool as it produces standardised SBML-qual format, supporting interoperability. It should be noted that some SBML-qual based tools and platforms such as BoolNet, CellNoptR, and CellCollective may require intermediate conversions in order to read and analyse the SBML-qual models. For instance, BoolNet and CellNoptR read only SBML-qual models generated by them but can't read the models generated by other tools. CellCollective platform can't read and analyse SBML-qual model generated by BoolNet and CellNoptR. In some cases, GINSIM tool alter or misrepresent Boolean functions. Therefore, it is necessary to manually adjust the BFs in this case.

Currently, a number of pathway-based models are analysed separately, while in fact they are interconnected. The integration of pathways is required to better understand the disease progression and therapeutic responses of PD. By integrating various signaling pathways, such as PI3K/AKT and Wnt, it is possible to identify crosstalk between these pathways and how they may contribute to the development and progression of PD. A broader investigation is necessary to fully understand the complex interplay between all pathways involved in PD. This information can inform the development of targeted therapeutic approaches and provide insight into potential targets for drug development.

5.7.2 Limitations of data type and integration

The stratification of the models relies on miRNAs. The miRNAs are small non-coding RNA molecules that play important roles in regulating gene expression and have been identified as potential biomarkers for a variety of diseases, including neurodegenerative disorders such as PD. However, some miRNAs that are frequently reported as biomarkers for these disorders are also found in other diseases, which limits their use. This is because these miRNAs may not be specific to a given disorder, but rather to a nonspecific processes like immune and inflammatory responses. To improve the accuracy and specificity of miRNA biomarkers for neurodegenerative disorders, other biomarkers could be involved, such as proteins or genes. In this work, the validation of miRNA biomarkers in independent studies is considered and can be further improved and updated to confirm their accuracy and specificity. The expression of key miRNAs in the PPMI study was compared to previously reported literature, and some miRNAs

are found to have mismatched expression levels (Table 4.7). Finally, it is essential to consider the specific context in which the miRNA is being expressed, as this may affect its regulation and potential use. Additionally, the level of an miRNA may be influenced by various factors such as the presence of other miRNAs, the availability of specific transcription factors, and the overall gene expression profile of the cell [350]. Further, the same miRNA may be differentially regulated in response to different drugs or drug combinations. This is because drugs can alter gene expression and miRNA regulation through various mechanisms, such as by inhibiting transcription factors or enzymes involved in miRNA processing. Therefore, it is important to carefully consider the context of experimental design when studying the role of miRNAs. By implementing and improving these strategies, it may be possible to improve the accuracy and specificity of miRNA biomarkers for neurodegenerative disorders and enhance their usefulness as diagnostic and prognostic tools.

The study uses a snapshot of dynamic data to simulate the mutation effect of T2DM on PD cohorts. This approach only captures one aspect of the comorbidity and may not accurately represent the complex relationship between PD and T2DM. A more comprehensive approach would be to analyze data that regularly tracks the progression of both disorders over time. This would provide a more accurate depiction of the comorbidity between PD and T2DM and allow for a deeper understanding of how these conditions interact and influence each other.

5.7.3 Limitations of the results applicability

Accuracy of simulation results depend on the input of accurate and relevant data, as well as the use of appropriate assumptions and algorithms. In this study, the simulation of some phenotypes requires further integration with omics data, allowing for precise staging during disease progression and better interpretations. A key question is whether the model is capable of exhibiting a healthy response to a disease as a defensive mechanism, or of showing no response to serious damage. In this context, the glycolysis phenotype needs to be adapted to differentiate between aerobic and anaerobic types. Additionally, the neuronal survival and apoptosis are the observable effects of the disease on the cells, such as the survival or death of the cells. These phenotypes are seen as high-level representations in the models because they are the visible outcomes of complex processes that occur at the molecular and cellular level. In order to fully understand the roles that neuronal survival and apoptosis play in disease progression, it is necessary to study these phenotypes in more detail, including the underlying molecular and cellular mechanisms.

Another limitation is that the molecular mechanism dynamics of a disease may vary from patient to patient, even within the same cohort. This can make it

difficult to accurately model the disease at the cohort level, as it may be difficult to account for the variations in molecular mechanism dynamics between patients. To improve the accuracy of the approach, it is necessary to study the disease at the individual patient level. Moreover, this work does not consider other factors that may affect the dynamics of the disease, such as gender and age.

Chapter 6

Conclusions

6.1 Key findings

In this study, BMs are automatically constructed from the PD map diagrams using the CaSQ and bipartite representations. These models are created in different formats, including SBML-qual and SIF to be analyzed with various tools. The accuracy of the models is critically evaluated through the analysis of their structural and dynamic properties.

Structural analysis of the models involved the calculation of basic topological features, such as the in/out degrees and interaction directionality. In/out degree refers to the number of incoming and outgoing connections a node (representing a biomolecule) has in the network. Interaction directionality refers to the direction in which the interaction between two biomolecules occurs. Further, centrality measures provide a way to evaluate the importance of a biomolecule in the model by considering its influence on the flow of information. For example, a biomolecule with a high betweenness centrality may act as a bridge between multiple biomolecules and therefore have a significant impact on the overall functioning of the model.

To verify the dynamic properties of the models and the underlying PD mechanisms, sensitivity analysis was performed involving small perturbations, such as knockouts (removal of a node from the network) and overexpressions (increased activity of a node). This analysis helped to ensure that the models accurately reflect the dynamic interactions and is robust against perturbations. The models were validated by comparing simulated and perturbed system-level behavior with validated behavior from the literature. The study used the TCA cycle, and the Wnt-PI3K/AKT pathway at various levels of complexity as examples.

Model stratification with PD cohort omics data was used to study the performance of BMs representing specific subgroups of patients. These BMs identified

different behaviour of studied molecular mechanisms during simulated disease progression. The models can also identify similarities of molecular activity across PD subtypes. This information can be used to support similarity-based differential diagnosis and to better understand the common crosstalks between different disease subtypes. This in turn may help to develop more targeted therapies that are more effective at managing common symptoms.

6.2 Implications of this study

The study shows that Boolean modelling is a promising tool for understanding the complexity of PD. By modelling the dynamic interactions between various biomolecules, researchers can gain insight into the underlying mechanisms of diseases and propose potential therapeutic strategies. BMs can be used to test hypotheses about the role of particular biomolecules in the development of a disease. Using a BM, the effect of different levels of a given biomolecule on the disease progression can be simulated and evaluated against experimental or observational data. In this way, BMs can be used to advance research hypotheses by providing a computational tool for testing and refining them based on the available data.

The study showed that the BMs of PD can also be used to explore the impact of multiple perturbations on the disease and identify patterns that may not be apparent from experimental or observational data alone. This can help to simulate and understand the effect of different combinations of therapeutics on the progression of a disease to propose the most promising therapeutic strategies.

Additionally, the study provides stratified BMs based on characteristics such as the severity of their disease, the presence of certain biomarkers, or the probability of their activity. Such stratification can help to improve the diagnosis of PD by analysing the specific factors that contribute to the disease progression. The models can be used to identify patterns and predict disease progression in different subgroups. This can be helpful in explaining the heterogeneity of a disease and developing diagnostic criteria. The models can also be used to tailor treatment strategies to the specific needs of each subgroup. This may involve using targeted interventions to slow the progression of the disease or providing more frequent follow-up care to monitor the effectiveness of these interventions.

6.3 Limitations

As previously discussed, BMs can be used to represent and analyze the relationships between different variables in a system or a process. However, like any

modelling approach, BMs are subject to certain limitations and risks. Construction of a model with insufficient details may lead to inaccurate predictions. Modellers may avoid this by using specific knowledge repositories of disease mechanisms, performing exploratory investigations and gathering information about the model from the literature and databases. It is possible to infer the missing details by integrating omics data, which identifies the missing components and optimizes the model's accuracy. In this regard, another potential risk is the incomplete data, e.g. due to issues with data collection or measurement, or lack of patients representing the entire disease spectrum. Incomplete data can introduce bias and lead to inaccurate conclusions. The incomplete data can affect the accuracy and reliability of the BMs by reducing the amount of information that is available for analysis.

BMs are based on assumptions and simplifications of the underlying biological processes involved in the disease. While these assumptions and simplifications can help to make the models more tractable and easier to analyze, they may not accurately reflect the complexity of the relationships between different factors specially with the incomplete data. If key biomolecules or phenotypes are missing, it may be difficult to accurately represent the relationships between them and draw conclusions about the underlying processes. This can lead to inaccuracies in the predictions made by the model, and ultimately limit its usefulness in guiding diagnosis and treatment decisions.

Model scale is also a significant challenge. Complex models are more difficult to analyze, and it is more difficult to reach attractor structures as the model scale increases [351, 352, 353]. To overcome this challenge, scientists propose different reduction approaches to control complex models. However, some reduction techniques, such as [117, 116], can be useful for simplifying and analyzing complex models, but it is important to keep in mind that they may not always accurately reflect the behavior of biological systems. Therefore, it is essential to carefully evaluate the limitations of these techniques and to verify the results of any analysis with experimental data.

6.4 Recommendations for the future research

BMs have shown promise as a tool for predicting disease mechanisms (Table 2.2) and for guiding the development of personalized treatments. However, there are several recommendations that should be considered in order to improve the use of BMs in clinical applications. One recommendation is to focus on improving data integration from various sources, such as gene expression data, protein-protein interaction data, and literature-based knowledge. This will require the development of robust and scalable methods for data integration, as well as the

establishment of standards for data representation and interoperability.

Cross-validation: Additional steps can be done to improve the evaluation of model performance by using cross-validation approaches. Cross-validation will involve dividing the omics data into training and test sets, training the model on the training set, and evaluating the model on the test set. By repeating this process multiple times using different portions of the data as the training and test sets, cross-validation allows for a more robust evaluation of the model's performance and can provide a more accurate estimate of its overall performance on the dataset.

Model precision: In order to accurately predict the behavior of biological systems, BMs should be able to capture the underlying mechanisms and dynamics of the system with a high degree of accuracy. In order to improve the accuracy of BMs, researchers must have access to high-quality data on the relationships between different factors that are believed to be involved and integrated in the models. It is crucial to carefully consider the potential for missing data during model construction and training and to utilize statistical techniques that can effectively check the model quality. By doing so, we can ensure that the BMs are thorough and comprehensive and can make useful predictions based on the available data.

In order to thoroughly understand and identify the effectiveness of therapeutic interventions, it is important to consider not just the calibrated initial states in a model, but also other possible initial states that may lead to the same probability of phenotype. This can help to provide a more comprehensive understanding of the various pathways and mechanisms involved in the disease process and may facilitate the identification of alternative therapeutic approaches. With regard to drug target interventions in Wnt-PI3K/AKT signaling, it may be useful to identify the minimum and maximum combinations of interventions that can alleviate disease conditions in order to create drug profiles and prioritize their application based on disease subtypes. This can help to optimize the therapeutic effectiveness of the interventions and potentially reduce the risk of undesirable side effects.

Standardization: In order to facilitate reproducibility and comparison of results across different studies, it is important to establish standards for model construction, validation, and evaluation. Addressing these recommendations will require a collaborative effort across the fields of mathematics, bioinformatics, and biology, as well as the development of systematic approaches for studying the modes of action and dosages of therapeutic interventions at a patient level. By following established protocols and best practices, it is possible to make the use

of BMs more understandable and reproducible, and to improve the development of decision-making pipelines based on these models in the future.

Scaling up mathematical models in systems biology - the role of HPC: In systems biology, the analysis of complex biological systems at the molecular level often requires the use of mathematical models such as BMs and ODEs. While ODEs are effective for small- to medium-scale systems, they have limitations when it comes to genome-scale analysis [354, 355]. In this context, BMs can offer a solution by simulation of interactions between various biomolecules on a larger scale, with thousands of nodes and edges. However, this increased complexity requires the use of larger machines with more computational power, such as HPC systems, distributed computing, specialized hardware, and optimization techniques.

As illustrated in Figure 6.1 from Kratochvil et al [356], the size of published models can greatly impact the computational demands of analysis. Even analyzing a relatively small number of model variants (4 million) using flux variability analysis can require more than 10^{18} operations, the threshold for human microbiome models. Models constructed from metagenomic data can also exceed this threshold quickly, highlighting the importance of efficient methods in the analysis of large models.

By successfully scaling up these models, modellers can make more accurate predictions about the response of diseases to different therapeutics and gain a deeper understanding of diseases at the molecular level [357]. Further research and development in this area has the potential to lead to even more powerful and accurate models in the future, making it a vital area of study in systems biology

6.5 Perspectives for interoperability and reproducibility of Boolean models

It is important to improve interoperability, annotations, and reproducibility of BMs.

- **Annotations** are important because they provide context and information about the models, which makes it easier for others to understand and use the models. BMs can be complex and difficult to interpret without additional information. By adding annotations to the models, a modeller can provide more context and clarity, which can help other researchers to understand the models and use them more effectively.
- **Interoperability** is important because it allows different models and tools to work together seamlessly. BMs can be used in combination with other

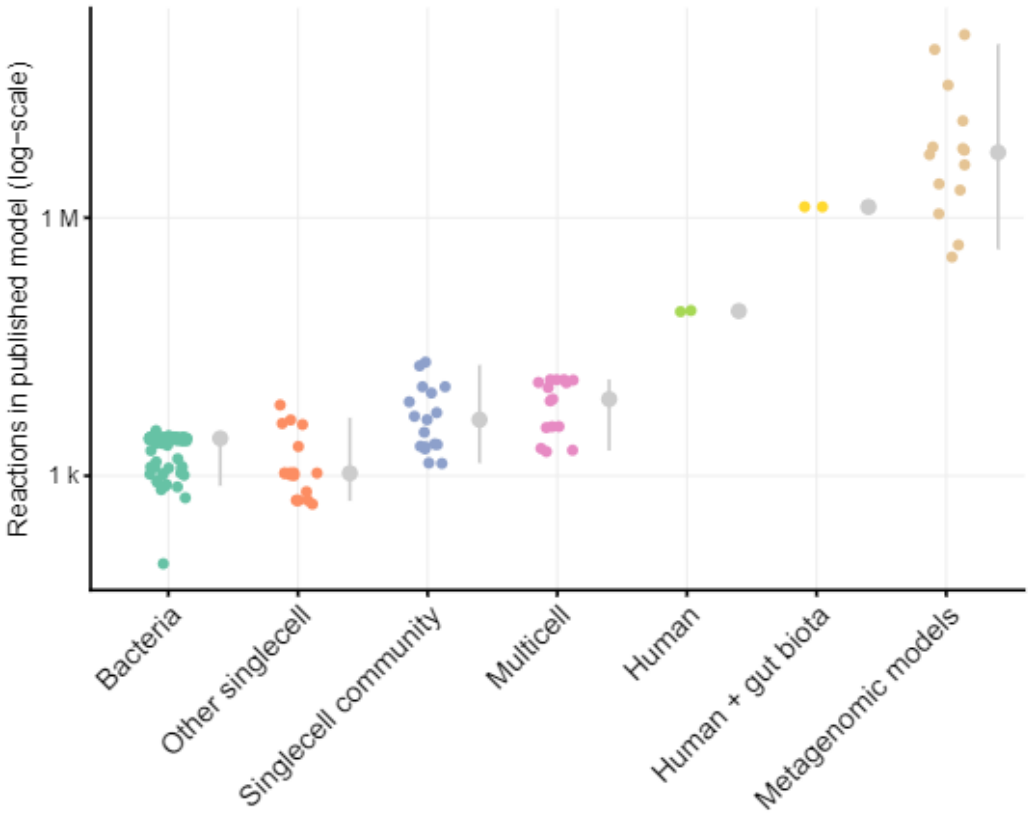


Figure 6.1 Examining the computational demands of large-scale modeling: model size and analysis efficiency. The size of published models can significantly impact the volume of computation required during analysis. Reproduced with permission from [356].

types of models and data sources to gain a more comprehensive understanding of biological systems. By improving interoperability, a researcher can ensure that models can be easily combined and analyzed in a way that maximizes their utility and effectiveness.

- **Reproducibility** is important because it allows others to validate and confirm the results of the model. BMs are often used to make predictions about the behavior of biological systems. By ensuring that models are reproducible, the results of the model can be independently verified and validated, which can help to increase confidence in the predictions made by the model.

Improvement of modelling standards: To improve the previous items, it is important to establish and follow best practices for model construction, validation, and evaluation. This can be achieved using standardised formats (e.g., SBMLpackages) which facilitate the development of logical modelling pipelines (e.g., ColoMoTo notebook). Repositories like GINsim and CellCollective allow users to create, annotate, and share models. Another important perspective is the integration of bioinformatic repositories with logical modelling. The BioModels platform (https://www.ebi.ac.uk/biomodels/) is one example of a repository that already supports logical models, but more work is needed to improve the interoperability of different systems and to make it easier for researchers to access and use the data and models available in these repositories.[5].

Model repositories: Maintaining model repositories and sharing models in easily interoperable formats is essential for improving reproducibility and advancing the field of computational biology. By providing a central location for storing and accessing models, model repositories enable researchers to build upon previous work and ensure that their own results are reliable and accurate. Model repositories also facilitate collaboration between researchers and promote transparency in research by providing a clear record of the models that have been developed and used in a given study. Examples of model repositories include BioModels and the Java Web Simulation repository of pathway/genome databases [358]. By utilizing these resources, researchers can access and reuse a wide range of models, enabling them to gain a greater understanding of biological systems and design more effective therapies for a range of diseases.

Modelling communities: The development of logical models in biology is a complex and multi-faceted process that requires the collaboration of various modelling communities, such as the Computational Modelling in Biology Network (COMBINE), Simulation Experiment Description Markup Language (SED-ML)

and SBML communities. These communities have recognized the need for best practices in the curation and annotation of logical models to facilitate their storage, reuse, and reproducibility. One notable example of this collaboration is the initiative of the ColoMoTo and SysMod communities to develop reproducibility scorecards, which provide a set of eight questions to evaluate the reusability and reproducibility of systems biology models [359]. This effort highlights the importance of automatic approaches to model annotation, quality assessment, and curation, as well as the need for a minimum amount of information to be defined and systematically applied in order to facilitate the storage and reuse of logical models [360].

In conclusion: As a member of the modelling community, I am committed to continuously strive for excellence and accuracy in this work. I believe that the importance of this work cannot be overstated, as accurate and reliable models are essential for driving progress in many areas of science and technology. To ensure that this work is reliable, relevant, and up-to-date, I make it a priority to stay abreast of the latest experimental workflows and research findings, and to adopt best practices for the modelling process. This may involve regularly updating and improving existing models, as well as developing dedicated approaches for automatically updating models as new data becomes available. I also believe that it is important to use a transparent and well-documented process in my modelling work, including clearly defining the goals of the model, the assumptions and limitations of the model, and the methods used to validate the model. This helps to ensure that the models are accurate, reliable, and relevant to the research question at hand. In addition, I make sure to use a diverse range of data sources and to carefully consider the quality and relevance of the data used to build and train the models. By using a diverse range of data, I can increase the applicability of the findings and reduce the risk of bias in the models. To conclude, I am dedicated to following best practices and continuously improving my work as a modeler. By doing so, I believe that this work can make significant contributions to the advancement of scientific knowledge and contribute to the progress in a wide range of fields.

References

[1] E.P.G.del Valle. *Disease networks and their contribution to disease understanding and drug repurposing: Evolution of the concept, techniques and data sources.* en. 2019.

[2] H. Ogata. "KEGG: Kyoto Encyclopedia of Genes and Genomes". es. In: *Nucleic Acids Res* 27 (1999), pp. 29–34.

[3] M. Wu. *A Dynamic Analysis of Insulin Signaling and Its Feedback Mechanisms: A Discrete Modeling Approach.* en. 2009.

[4] D.N. Slenter. "WikiPathways: a multifaceted pathway database bridging metabolomics to other omics research". ro. In: *Nucleic Acids Res* 46 (2018), pp. 661–667.

[5] R.S. Malik-Sheriff. "BioModels—15 years of sharing computational models in life science". en. In: *Nucleic Acids Res* 48 (2020), pp. 407–415.

[6] F. Kazantsev. "MAMMOTh: A new database for curated mathematical models of biomolecular systems". pt. In: *J. Bioinform. Comput. Biol* 16 (2017), p. 1740010.

[7] T. Helikar. "The Cell Collective: Toward an open and collaborative approach to systems biology". en. In: *BMC Syst. Biol* 6 (2012), p. 96.

[8] A. Naldi. "Logical Modeling and Analysis of Cellular Regulatory Networks With GINsim 3.0". en. In: *Front. Physiol* 9 (2018).

[9] U. Wittig. "SABIO-RK–database for biochemical reaction kinetics". en. In: *Nucleic Acids Res* 40 (2012), pp. 790–796.

[10] A. Mazein. "Systems medicine disease maps: community-driven comprehensive representation of disease mechanisms". fr. In: *NPJ Syst. Biol. Appl* 4 (2018), p. 21.

[11] A. Funahashi. "CellDesigner: A Graphical Biological Network Editor and Workbench Interfacing Simulator". en. In: *Introduction to Systems Biology*. Ed. by Choi In and S. Totowa, NJ: Humana Press, 2007, pp. 422–434.

[12] H. Balci. "Newt: a comprehensive web-based tool for viewing, constructing and analyzing biological maps". en. In: *Bioinformatics* (2020).

[13] R. Wiese. "yFiles — Visualization and Automatic Layout of Graphs". en. In: *Graph Drawing Software, Mathematics and Visualization*. Ed. by Jünger In, M., and P. Mutzel. Berlin, Heidelberg: Springer, 2004, pp. 173–191.

[14] I. Kuperstein. “NaviCell: a web-based environment for navigation, curation and maintenance of large molecular interaction maps”. en. In: *BMC Syst. Biol* 7 (2013), p. 100.

[15] P. Gawron. “MINERVA-a platform for visualization and curation of molecular interaction networks”. en. In: *NPJ Syst. Biol. Appl* 2 (2016), p. 16020.

[16] S.S. Aghamiri. *Automated inference of Boolean models from molecular interaction maps using CaSQ*. en. Bioinformatics.

[17] D. Silk. “Model Selection in Systems Biology Depends on Experimental Design”. es. In: *PLOS Comput. Biol* 10 (2014), p. 1003650.

[18] E. Dubrova. “Random Multiple-Valued Networks: Theory and Applications”. en. In: *36th International Symposium on Multiple-Valued Logic (ISMVL’06)*. 2006, pp. 27–27.

[19] W.M.P. Aalst. “Petri Nets”. ro. In: *Encyclopedia of Database Systems*. Ed. by L. In LIU and M.T. ÖZSU. Boston, MA: Springer US, 2009, pp. 2103–2108.

[20] W. Walter. *Ordinary Differential Equations Springer-Verlag*. de. Place: New York. 1998.

[21] M. Ilea, M. Turnea, and M. Rotariu. “Ordinary differential equations with applications in molecular biology”. eng. In: *Revista Medico-Chirurgicala a Societatii De Medici Si Naturalisti Din Iasi* 116.1 (2012), pp. 347–352. ISSN: 0048-7848.

[22] Dale Purves et al. *Neurosciences*. fr. Google-Books-ID: JDSZDwAAQBAJ. De Boeck Superieur, June 2019. ISBN: 978-2-8073-1492-4.

[23] Apoorva Gopalakrishna and Sheila A. Alexander. “Understanding Parkinson Disease: A Complex and Multifaceted Illness”. eng. In: *The Journal of Neuroscience Nursing: Journal of the American Association of Neuroscience Nurses* 47.6 (Dec. 2015), pp. 320–326. ISSN: 1945-2810. DOI: 10.1097/JNN.0000000000000162.

[24] Jose A. Santiago, Virginie Bottero, and Judith A. Potashkin. “Biological and Clinical Implications of Comorbidities in Parkinson’s Disease”. In: *Frontiers in Aging Neuroscience* 9 (Dec. 2017), p. 394. ISSN: 1663-4365. DOI: 10.3389/fnagi.2017.00394. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5722846/ (visited on 12/03/2022).

[25] Kevin McFarthing, Danielle Larson, and Tanya Simuni. “Clinical Trial Highlights – GLP-1 agonists”. In: *Journal of Parkinson’s Disease* 10.2 (), pp. 355–368. ISSN: 1877-7171. DOI: 10.3233/JPD-200002. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7242848/ (visited on 12/03/2022).

[26] Sophie L. Farrow, Antony A. Cooper, and Justin M. O’Sullivan. “Redefining the hypotheses driving Parkinson’s diseases research”. en. In: *npj Parkinson’s Disease* 8.1 (Apr. 2022). Number: 1 Publisher: Nature Publishing Group, pp. 1–7. ISSN: 2373-8057. DOI: 10.1038/s41531-022-00307-w. URL: https://www.nature.com/articles/s41531-022-00307-w (visited on 12/08/2022).

[27] Debmalya Barh et al. “In silico disease model: from simple networks to complex diseases”. In: *Animal Biotechnology* (2020), pp. 441–460. DOI: 10.1016/B978-0-12-811710-1.00020-3. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7325851/ (visited on 12/08/2022).

[28] Elaina M. Maldonado et al. “Integration of Genome Scale Metabolic Networks and Gene Regulation of Metabolic Enzymes With Physiologically Based Pharmacokinetics”. In: *CPT: Pharmacometrics & Systems Pharmacology* 6.11 (Nov. 2017), pp. 732–746. ISSN: 2163-8306. DOI: 10.1002/psp4.12230. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5702902/ (visited on 12/19/2022).

[29] R. Albert and J. Thakar. “Boolean modeling: a logic-based dynamic approach for understanding signaling and regulatory networks and for making useful predictions”. en. In: *Wiley Interdiscip. Rev. Syst. Biol. Med* 6 (2014), pp. 353–369.

[30] R. Albert and J. Thakar. “Boolean modeling: a logic-based dynamic approach for understanding signaling and regulatory networks and for making useful predictions”. en. In: *Wiley Interdiscip. Rev. Syst. Biol. Med* 6 (2014), pp. 353–369.

[31] C. Terfve. “CellNOptR: a flexible toolkit to train protein signaling networks to data using multiple logic formalisms”. en. In: *BMC Syst. Biol* 6 (2012), p. 133.

[32] P. Bloomingdale. “Boolean network modeling in systems pharmacology”. nl. In: *J. Pharmacokinet. Pharmacodyn* 45 (2018), pp. 159–180.

[33] F. Eduati. “Patient-specific logic models of signaling pathways from screenings on cancer biopsies to prioritize personalized combination therapies”. en. In: *Mol. Syst. Biol* 16 (2020), p. 8664.

[34] R.-S. Wang. “Boolean modeling in systems biology: an overview of methodology and applications”. en. In: *Phys. Biol* 9 (2012), p. 055001.

[35] Å. Flobak. “Discovery of Drug Synergies in Gastric Cancer Cells Predicted by Logical Modeling”. en. In: *PLOS Comput. Biol* 11 (2015), p. 1004426.

[36] Y.-K. Kwon. “Dynamical Robustness Against Multiple Mutations in Signaling Networks”. en. In: *IEEEACM Trans. Comput. Biol. Bioinforma. IEEE ACM* 13 (2015).

[37] Jonas Béal et al. “Personalized logical models to investigate cancer response to BRAF treatments in melanomas and colorectal cancers”. en. In: *PLOS Computational Biology* 17.1 (Jan. 2021). Publisher: Public Library of Science, e1007900. ISSN: 1553-7358. DOI: 10.1371/journal.pcbi.1007900. URL: https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1007900 (visited on 11/27/2022).

[38] A. Poret and C. Guziolowski. “Therapeutic target discovery using Boolean network attractors: improvements of kali”. en. In: *R. Soc. Open Sci* 5 (2018).

[39] L. Siegle. “A Boolean network of the crosstalk between IGF and Wnt signaling in aging satellite cells”. en. In: *PloS One* 13 (2018), p. 0195126.

[40] C. Campbell and R. Albert. “Stabilization of perturbed Boolean network attractors through compensatory interactions”. en. In: *BMC Syst. Biol* 8 (2014), p. 53.

[41] E. Gjerga. *Converting networks to predictive logic models from perturbation signalling data with CellNOpt Bioinformatics*. en. 2020.

[42] I. Albert. “Boolean network simulations for life scientists”. en. In: *Source Code Biol. Med* 3 (2008), p. 16.

[43] V. Tran. “On the underlying assumptions of threshold Boolean networks as a model for genetic regulatory network behavior”. en. In: *Front. Genet* 4 (2013).

[44] Bianca Marchetti et al. “Parkinson’s disease, aging and adult neurogenesis: Wnt/β-catenin signalling as the key to unlock the mystery of endogenous brain repair”. In: *Aging Cell* 19.3 (Mar. 2020), e13101. ISSN: 1474-9718. DOI: 10.1111/acel.13101. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7059166/ (visited on 12/26/2022).

[45] Johannes Meiser, Daniel Weindl, and Karsten Hiller. “Complexity of dopamine metabolism”. In: *Cell Communication and Signaling* 11.1 (May 2013), p. 34. ISSN: 1478-811X. DOI: 10.1186/1478-811X-11-34. URL: https://doi.org/10.1186/1478-811X-11-34 (visited on 11/23/2022).

[46] Amy M. Smith et al. “Mitochondrial dysfunction and increased glycolysis in prodromal and early Parkinson’s blood cells”. eng. In: *Movement Disorders: Official Journal of the Movement Disorder Society* 33.10 (Oct. 2018), pp. 1580–1590. ISSN: 1531-8257. DOI: 10.1002/mds.104.

[47] Leonie M. Heger et al. “Mitochondrial Phenotypes in Parkinson’s Diseases—A Focus on Human iPSC-Derived Dopaminergic Neurons”. en. In: *Cells* 10.12 (Dec. 2021). Number: 12 Publisher: Multidisciplinary Digital Publishing Institute, p. 3436. ISSN: 2073-4409. DOI: 10.3390/cells10123436. URL: https://www.mdpi.com/2073-4409/10/12/3436 (visited on 11/23/2022).

[48] Yingqing Wang et al. “Beneficial effects of PGC-1α in the substantia nigra of a mouse model of MPTP-induced dopaminergic neurotoxicity”. In: *Aging (Albany NY)* 11.20 (Oct. 2019), pp. 8937–8950. ISSN: 1945-4589. DOI: 10.18632/aging.102357. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6834419/ (visited on 12/10/2022).

[49] Nadia Turton et al. “Coenzyme Q_{10} and the exclusive club of diseases that show a limited response to treatment”. en. In: *Expert Opinion on Orphan Drugs* 9.5 (May 2021), pp. 151–160. ISSN: 2167-8707. DOI: 10.1080/21678707.2021.1932459. URL: https://www.tandfonline.com/doi/full/10.1080/21678707.2021.1932459 (visited on 12/10/2022).

[50] Mayumi Takahashi and Kazuhide Takahashi. “Water-soluble CoQ10 as A Promising Anti-aging Agent for Neurological Dysfunction in Brain Mitochondria”. In: *Antioxidants* 8.3 (Mar. 2019), p. 61. ISSN: 2076-3921. DOI: 10.3390/antiox8030061. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6466529/ (visited on 12/10/2022).

[51] Sajad Fakhri et al. “Targeting Mitochondria by Plant Secondary Metabolites: A Promising Strategy in Combating Parkinson’s Disease”. en. In: *International Journal of Molecular Sciences* 22.22 (Jan. 2021). Number: 22 Publisher: Multidisciplinary Digital Publishing Institute, p. 12570. ISSN: 1422-0067. DOI: 10.3390/ijms222212570. URL: https://www.mdpi.com/1422-0067/22/22/12570 (visited on 12/10/2022).

[52] Alexander Grotemeyer et al. “Neuroinflammation in Parkinson’s Disease – Putative Pathomechanisms and Targets for Disease-Modification”. In: *Frontiers in Immunology* 13 (2022). ISSN: 1664-3224. URL: https://www.frontiersin.org/articles/10.3389/fimmu.2022.878771 (visited on 12/11/2022).

[53] Anumantha G. Kanthasamy et al. “Environmental neurotoxin dieldrin induces apoptosis via caspase-3-dependent proteolytic activation of protein kinase C delta (PKCdelta): Implications for neurodegeneration in Parkinson’s disease”. eng. In: *Molecular Brain* 1 (Oct. 2008), p. 12. ISSN: 1756-6606. DOI: 10.1186/1756-6606-1-12.

[54] E. Srinivasan et al. “Alpha-Synuclein Aggregation in Parkinson’s Disease”. In: *Frontiers in Medicine* 8 (Oct. 2021), p. 736978. ISSN: 2296-858X. DOI: 10.3389/fmed.2021.736978. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8558257/ (visited on 12/19/2022).

[55] Maxime Teixeira et al. “Alpha-Synuclein and the Endolysosomal System in Parkinson’s Disease: Guilty by Association”. en. In: *Biomolecules* 11.9 (Sept. 2021). Number: 9 Publisher: Multidisciplinary Digital Publishing Institute, p. 1333. ISSN: 2218-273X. DOI: 10.3390/biom11091333. URL: https://www.mdpi.com/2218-273X/11/9/1333 (visited on 12/19/2022).

[56] Darius Ebrahimi-Fakhari, Lara Wahlster, and Pamela J. McLean. “Protein degradation pathways in Parkinson’s disease: curse or blessing”. eng. In: *Acta Neuropathologica* 124.2 (Aug. 2012), pp. 153–172. ISSN: 1432-0533. DOI: 10.1007/s00401-012-1004-6.

[57] Tapan Behl et al. “Exploring the Role of Ubiquitin–Proteasome System in Parkinson’s Disease”. en. In: *Molecular Neurobiology* 59.7 (July 2022), pp. 4257–4273. ISSN: 1559-1182. DOI: 10.1007/s12035-022-02851-1. URL: https://doi.org/10.1007/s12035-022-02851-1 (visited on 12/19/2022).

[58] Shahrokh Lorzadeh et al. “Autophagy and the Wnt signaling pathway: A focus on Wnt/β-catenin signaling”. eng. In: *Biochimica Et Biophysica Acta. Molecular Cell Research* 1868.3 (Mar. 2021), p. 118926. ISSN: 1879-2596. DOI: 10.1016/j.bbamcr.2020.118926.

[59] B. Post et al. “Multidisciplinary care for Parkinson’s disease: not if, but how!” eng. In: *Practical Neurology* 11.2 (Apr. 2011), pp. 58–61. ISSN: 1474-7766. DOI: 10.1136/jnnp.2011.241604.

[60] Karen Nuytemans et al. “Genetic etiology of Parkinson disease associated with mutations in the SNCA, PARK2, PINK1, PARK7, and LRRK2 genes: a mutation update”. eng. In: *Human Mutation* 31.7 (July 2010), pp. 763–780. ISSN: 1098-1004. DOI: 10.1002/humu.21277.

[61] Dilan Athauda et al. “The Impact of Type 2 Diabetes in Parkinson’s Disease”. en. In: *Movement Disorders* 37.8 (2022). _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/mds.29122, pp. 1612–1623. ISSN: 1531-8257. DOI: 10.1002/mds.29122. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/mds.29122 (visited on 12/10/2022).

[62] Haiyang Yu et al. “Association between Parkinson’s Disease and Diabetes Mellitus: From Epidemiology, Pathophysiology and Prevention to Treatment”. eng. In: *Aging and Disease* 13.6 (Dec. 2022), pp. 1591–1605. ISSN: 2152-5250. DOI: 10.14336/AD.2022.0325.

[63] Abdallah Hassan et al. “Diabetes Mellitus and Parkinson’s Disease: Shared Pathophysiological Links and Possible Therapeutic Implications”. eng. In: *Cureus* 12.8 (Aug. 2020), e9853. ISSN: 2168-8184. DOI: 10.7759/cureus.9853.

[64] Jill L. Smith et al. “Levodopa with carbidopa diminishes glycogen concentration, glycogen synthase activity, and insulin-stimulated glucose transport in rat skeletal muscle”. eng. In: *Journal of Applied Physiology (Bethesda, Md.: 1985)* 97.6 (Dec. 2004), pp. 2339–2346. ISSN: 8750-7587. DOI: 10.1152/japplphysiol.01219.2003.

[65] Liying Yang et al. “The Role of Insulin/IGF-1/PI3K/Akt/GSK3β Signaling in Parkinson’s Disease Dementia”. In: *Frontiers in Neuroscience* 12 (Feb. 2018), p. 73. ISSN: 1662-4548. DOI: 10.3389/fnins.2018.00073. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5826217/ (visited on 11/04/2021).

[66] Tomasz Chmiela et al. “If Not Insulin Resistance so What? - Comparison of Fasting Glycemia in Idiopathic Parkinson’s Disease and Atypical Parkinsonism”. eng. In: *Diabetes, Metabolic Syndrome and Obesity: Targets and Therapy* 15 (2022), pp. 1451–1460. ISSN: 1178-7007. DOI: 10.2147/DMSO.S359856.

[67] Elliot Hogg et al. “High Prevalence of Undiagnosed Insulin Resistance in Non-Diabetic Subjects with Parkinson’s Disease”. eng. In: *Journal of Parkinson’s Disease* 8.2 (2018), pp. 259–265. ISSN: 1877-718X. DOI: 10.3233/JPD-181305.

[68] Carolina Sportelli et al. “Metformin as a Potential Neuroprotective Agent in Prodromal Parkinson’s Disease-Viewpoint”. eng. In: *Frontiers in Neurology* 11 (2020), p. 556. ISSN: 1664-2295. DOI: 10.3389/fneur.2020.00556.

[69] Kenneth Marek et al. “The Parkinson’s progression markers initiative (PPMI) – establishing a PD biomarker cohort”. In: *Annals of Clinical and Translational Neurology* 5.12 (Oct. 2018), pp. 1460–1477. ISSN: 2328-9503. DOI: 10.1002/acn3.644. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6292383/ (visited on 12/10/2022).

[70] Fabian Kern et al. *Deep sncRNA-seq of the PPMI cohort to study Parkinson’s disease progression*. en. Pages: 2020.06.01.127092 Section: New Results. June 2020. DOI: 10.1101/2020.06.01.127092. URL: https://www.biorxiv.org/content/10.1101/2020.06.01.127092v1 (visited on 11/23/2022).

[71] Yohei Mukai and Miho Murata. “[Scans without Evidence of Dopamine Deficit (SWEDDs)]”. jpn. In: *Brain and Nerve = Shinkei Kenkyu No Shinpo* 68.1 (Jan. 2016), pp. 73–81. ISSN: 1881-6096. DOI: 10.11477/mf.1416200350.

[72] Philipp Mahlknecht, Klaus Seppi, and Werner Poewe. “The Concept of Prodromal Parkinson’s Disease”. eng. In: *Journal of Parkinson’s Disease* 5.4 (2015), pp. 681–697. ISSN: 1877-718X. DOI: 10.3233/JPD-150685.

[73] Geraldine Hipp et al. “The Luxembourg Parkinson’s Study: A Comprehensive Approach for Stratification and Early Diagnosis”. In: *Frontiers in Aging Neuroscience* 10 (Oct. 2018), p. 326. ISSN: 1663-4365. DOI: 10.3389/fnagi.2018.00326. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6216083/ (visited on 12/19/2022).

[74] Michael Lawton et al. “Parkinson’s Disease Subtypes in the Oxford Parkinson Disease Centre (OPDC) Discovery Cohort”. eng. In: *Journal of Parkinson’s Disease* 5.2 (2015), pp. 269–279. ISSN: 1877-718X. DOI: 10.3233/JPD-140523.

[75] D. Silk. “Model Selection in Systems Biology Depends on Experimental Design”. es. In: *PLOS Comput. Biol* 10 (2014), p. 1003650.

[76] Jingwen Yan et al. “Network approaches to systems biology analysis of complex disease: integrative methods for multi-omics data”. In: *Briefings in Bioinformatics* 19.6 (June 2017), pp. 1370–1381. ISSN: 1467-5463. DOI: 10.1093/bib/bbx066. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6454489/ (visited on 12/19/2022).

[77] Ahmed Abdelmonem Hemedan et al. “Boolean modelling as a logic-based dynamic approach in systems medicine”. en. In: *Computational and Structural Biotechnology Journal* 20 (Jan. 2022), pp. 3161–3172. ISSN: 2001-0370. DOI: 10.1016/j.csbj.2022.06.035. URL: https://www.sciencedirect.com/science/article/pii/S2001037022002495 (visited on 12/19/2022).

[78] W.M.P. Aalst. “Petri Nets”. ro. In: *Encyclopedia of Database Systems*. Ed. by L. In LIU and M.T. ÖZSU. Boston, MA: Springer US, 2009, pp. 2103–2108.

[79] W. Walter. *Ordinary Differential Equations Springer-Verlag*. de. New York, 1998.

[80] E. Dubrova. “Random Multiple-Valued Networks: Theory and Applications”. en. In: *36th International Symposium on Multiple-Valued Logic (ISMVL'06)*. 2006, pp. 27–27.

[81] M. Ilea. “Ordinary differential equations with applications in molecular biology”. en. In: *Rev. Med. Chir. Soc. Med. Nat. Iasi* 116 (2012), pp. 347–352.

[82] T. Schlitt and A. Brazma. “Current approaches to gene regulatory network modelling”. en. In: *BMC Bioinformatics* 8 (2007), p. 9.

[83] Monika Heiner, David Gilbert, and Robin Donaldson. “Petri Nets for Systems and Synthetic Biology”. en. In: *Formal Methods for Computational Systems Biology*. Ed. by Marco Bernardo, Pierpaolo Degano, and Gianluigi Zavattaro. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, 2008, pp. 215–264. ISBN: 978-3-540-68894-5. DOI: 10.1007/978-3-540-68894-5_7.

[84] Christoph Brinkrolf and Lennart Ochel. “Comprehensive Open-Source Petri Net Toolchain for Modeling and Simulation in Systems Biology”. en. In: *Integrative Bioinformatics: History and Future*. Ed. by Ming Chen and Ralf Hofestädt. Singapore: Springer, 2022, pp. 269–289. ISBN: 9789811667954. DOI: 10.1007/978-981-16-6795-4_13. URL: https://doi.org/10.1007/978-981-16-6795-4_13 (visited on 12/19/2022).

[85] Clemens Kreutz. “A New Approximation Approach for Transient Differential Equation Models”. In: *Frontiers in Physics* 8 (2020). ISSN: 2296-424X. URL: https://www.frontiersin.org/articles/10.3389/fphy.2020.00070 (visited on 12/19/2022).

[86] Jonathan R. Karr et al. “A whole-cell computational model predicts phenotype from genotype”. eng. In: *Cell* 150.2 (July 2012), pp. 389–401. ISSN: 1097-4172. DOI: 10.1016/j.cell.2012.05.044.

[87] Markus W. Covert et al. “Integrating metabolic, transcriptional regulatory and signal transduction models in Escherichia coli”. eng. In: *Bioinformatics (Oxford, England)* 24.18 (Sept. 2008), pp. 2044–2050. ISSN: 1367-4811. DOI: 10.1093/bioinformatics/btn352.

[88] C. Müssel. “BiTrinA—multiscale binarization and trinarization with quality analysis”. en. In: *Bioinformatics* 32 (2016), pp. 465–468.

[89] A. Karanam. *BoolSim, a Graphical Interface for Open Access Boolean Network Simulations and Use in Guard Cell CO2 Signaling*. en. bioRxiv, 2021.

[90] A. Di Cara. “Dynamic simulation of regulatory networks using SQUAD”. et. In: *BMC Bioinformatics* 8 (2007), p. 462.

[91] F. Büchel. “Path2Models: large-scale generation of computational models from biochemical pathway maps”. en. In: *BMC Syst. Biol* 7 (2013), p. 116.

[92] J.C. Romers and M. Krantz. *rxncon 2.0: a language for executable molecular systems biology*. pt. 2017.

[93] S.S. Aghamiri. *Automated inference of Boolean models from molecular interaction maps using CaSQ*. en. Bioinformatics.

[94] R.-S. Wang. "Boolean modeling in systems biology: an overview of methodology and applications". en. In: *Phys. Biol* 9 (2012), p. 055001.

[95] B. Drossel. *Random Boolean Networks. In, Reviews of Nonlinear Dynamics and Complexity*. en. John Wiley & Sons, Ltd, 2008, pp. 69–110.

[96] S. Kauffman. "Genetic networks with canalyzing Boolean rules are always stable". en. In: *Proc. Natl. Acad. Sci* 101 (2004), pp. 17102–17107.

[97] S. Bilke and F. Sjunnesson. "Stability of the Kauffman model". en. In: *Phys. Rev. E Stat. Nonlin. Soft Matter Phys* 65 (2002), p. 016129.

[98] A. Garg. "Synchronous versus asynchronous modeling of gene regulatory networks". en. In: *Bioinformatics* 24 (2008), pp. 1917–1925.

[99] M. Irwin and Z. Wang. "Dynamic Systems Modeling". en. In: *In, The International Encyclopedia of Communication Research Methods*. American Cancer Society, 2017, pp. 1–12.

[100] C. Campbell and R. Albert. "Stabilization of perturbed Boolean network attractors through compensatory interactions". en. In: *BMC Syst. Biol* 8 (2014), p. 53.

[101] J.D. Schwab. "Concepts in Boolean network modeling: What do they all mean? Comput". en. In: *Struct. Biotechnol. J* 18 (2020), pp. 571–582.

[102] C. Müssel. "BoolNet–an R package for generation, reconstruction and analysis of Boolean networks". en. In: *Bioinforma. Oxf. Engl* 26 (2010), pp. 1378–1380.

[103] M. Chaves. *Boolean dynamics revisited through feedback interconnections*. en. Nat. Comput, 2018.

[104] M. Grieb. "Predicting Variabilities in Cardiac Gene Expression with a Boolean Network Incorporating Uncertainty". en. In: *PLOS ONE* 10 (2015), p. 0131832.

[105] J.D. Schwab. "Concepts in Boolean network modeling: What do they all mean? Comput". en. In: *Struct. Biotechnol. J* 18 (2020), pp. 571–582.

[106] M. Hopfensitz. "Multiscale binarization of gene expression data for reconstructing Boolean networks". en. In: *IEEE/ACM Trans. Comput. Biol. Bioinform* 9 (2012), pp. 487–498.

[107] K. Klemm and S. Bornholdt. "Stable and unstable attractors in Boolean networks". en. In: *Phys. Rev. E* 72 (2005), p. 055101.

[108] A. Fauré. "Dynamical analysis of a generic Boolean model for the control of the mammalian cell cycle". it. In: *Bioinforma. Oxf. Engl* 22 (2006), pp. 124–131.

[109] I. Shmulevich. "The role of certain Post classes in Boolean network models of genetic networks". en. In: *Proc. Natl. Acad. Sci. U. S. A* 100 (2003), pp. 10734–10739.

[110] Armin Biere. “PicoSAT Essentials”. en. In: *Journal on Satisfiability, Boolean Modeling and Computation* 4.2-4 (Jan. 2008). Publisher: IOS Press, pp. 75–97. DOI: 10.3233/SAT190039. URL: https://content.iospress.com/articles/journal-on-satisfiability-boolean-modeling-and-computation/sat190039 (visited on 11/24/2022).

[111] Colin Campbell and Réka Albert. “Edgetic perturbations to eliminate fixed-point attractors in Boolean regulatory networks”. In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 29.2 (Feb. 2019), p. 023130. ISSN: 1054-1500. DOI: 10.1063/1.5083060. URL: https://aip.scitation.org/doi/abs/10.1063/1.5083060 (visited on 07/04/2019).

[112] A. Saadatpour. “A Reduction Method for Boolean Network Models Proven to Conserve Attractors”. en. In: *SIAM J. Appl. Dyn. Syst* 12 (2013), pp. 1997–2011.

[113] A. Veliz-Cuba. “Reduction of Boolean network models”. en. In: *J. Theor. Biol* 289 (2011), pp. 167–172.

[114] A. Veliz-Cuba. “Steady state analysis of Boolean molecular network models via model reduction and computational algebra”. en. In: *BMC Bioinformatics* 15 (2014), p. 221.

[115] A. Veliz-Cuba. “Reduction of Boolean network models”. en. In: *J. Theor. Biol* 289 (2011), pp. 167–172.

[116] A. Veliz-Cuba. “Steady state analysis of Boolean molecular network models via model reduction and computational algebra”. en. In: *BMC Bioinformatics* 15 (2014), p. 221.

[117] M. Moradi. “A Boolean network control algorithm guided by forward dynamic programming”. en. In: *PLOS ONE* 14 (2019), p. 0215449.

[118] Loïc Paulevé et al. “Reconciling qualitative, abstract, and scalable modeling of biological networks”. en. In: *Nature Communications* 11.1 (Aug. 2020), p. 4256. ISSN: 2041-1723. DOI: 10.1038/s41467-020-18112-5. URL: https://www.nature.com/articles/s41467-020-18112-5 (visited on 02/07/2022).

[119] Loïc Paulevé and Adrien Richard. “Static Analysis of Boolean Networks Based on Interaction Graphs: A Survey”. en. In: *Electronic Notes in Theoretical Computer Science*. Proceedings of the 2nd International Workshop on Static Analysis and Systems Biology (SASB 2011) 284 (June 2012), pp. 93–104. ISSN: 1571-0661. DOI: 10.1016/j.entcs.2012.05.017. URL: https://www.sciencedirect.com/science/article/pii/S1571066112000187 (visited on 12/19/2022).

[120] Ivan Rodrigo Wolf, Rafael Plana Simões, and Guilherme Targino Valente. “Three topological features of regulatory networks control life-essential and specialized subsystems”. en. In: *Scientific Reports* 11.1 (Dec. 2021). Number: 1 Publisher: Nature Publishing Group, p. 24209. ISSN: 2045-2322. DOI: 10.1038/s41598-021-03625-w. URL: https://www.nature.com/articles/s41598-021-03625-w (visited on 12/19/2022).

[121] Yufei Xiao and Edward R. Dougherty. "The impact of function perturbations in Boolean networks". In: *Bioinformatics* 23.10 (May 2007), pp. 1265–1273. ISSN: 1367-4803. DOI: 10.1093/bioinformatics/btm093. URL: https://doi.org/10.1093/bioinformatics/btm093 (visited on 02/07/2022).

[122] H.-C. Trinh and Y.-K. Kwon. "RMut: R package for a Boolean sensitivity analysis against various types of mutations". en. In: *PLOS ONE* 14 (2019), p. 0213736.

[123] T. Chatain. *Concurrency in Boolean networks*. en. Nat. Comput, 2019.

[124] K. Kobayashi and K. Hiraishi. "Optimization-Based Approaches to Control of Probabilistic Boolean Networks". en. In: *Algorithms* 10 (2017), p. 31.

[125] A. Poret and C. Guziolowski. "Therapeutic target discovery using Boolean network attractors: improvements of kali". en. In: *R. Soc. Open Sci* 5 (2018).

[126] M. Hu. "An efficient algorithm to identify the optimal one-bit perturbation based on the basin-of-state size of Boolean networks". en. In: *Sci. Rep* 6 (2016), p. 26247.

[127] H.-C. Trinh and Y.-K. Kwon. "RMut: R package for a Boolean sensitivity analysis against various types of mutations". en. In: *PLOS ONE* 14 (2019), p. 0213736.

[128] Paul Shannon et al. "Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks". en. In: *Genome Research* 13.11 (Nov. 2003). Company: Cold Spring Harbor Laboratory Press Distributor: Cold Spring Harbor Laboratory Press Institution: Cold Spring Harbor Laboratory Press Label: Cold Spring Harbor Laboratory Press Publisher: Cold Spring Harbor Lab, pp. 2498–2504. ISSN: 1088-9051, 1549-5469. DOI: 10.1101/gr.1239303. URL: https://genome.cshlp.org/content/13/11/2498 (visited on 12/19/2022).

[129] Dénes Türei, Tamás Korcsmáros, and Julio Saez-Rodriguez. "OmniPath: guidelines and gateway for literature-curated signaling pathway resources". en. In: *Nature Methods* 13.12 (Dec. 2016). Number: 12 Publisher: Nature Publishing Group, pp. 966–967. ISSN: 1548-7105. DOI: 10.1038/nmeth.4077. URL: https://www.nature.com/articles/nmeth.4077 (visited on 12/19/2022).

[130] Luana Licata et al. "SIGNOR 2.0, the SIGnaling Network Open Resource 2.0: 2019 update". In: *Nucleic Acids Research* 48.D1 (Jan. 2020), pp. D504–D510. ISSN: 0305-1048. DOI: 10.1093/nar/gkz949. URL: https://doi.org/10.1093/nar/gkz949 (visited on 12/19/2022).

[131] A. Karanam. *BoolSim, a Graphical Interface for Open Access Boolean Network Simulations and Use in Guard Cell CO2 Signaling*. en. bioRxiv, 2021.

[132] A. Di Cara. "Dynamic simulation of regulatory networks using SQUAD". et. In: *BMC Bioinformatics* 8 (2007), p. 462.

[133] G. Batt. "Genetic Network Analyzer: A Tool for the Qualitative Modeling and Simulation of Bacterial Regulatory Networks". en. In: *Bacterial Molecular Networks: Methods and Protocols, Methods in Molecular Biology*. Ed. by van Helden In and J. New York, NY: Springer, 2012, pp. 439–462.

[134] Aurélien Naldi et al. “Cooperative development of logical modelling standards and tools with CoLoMoTo”. In: *Bioinformatics* 31.7 (Apr. 2015), pp. 1154–1159. ISSN: 1367-4803. DOI: 10.1093/bioinformatics/btv013. URL: https://doi.org/10.1093/bioinformatics/btv013 (visited on 12/19/2022).

[135] Michael Hucka et al. “The Systems Biology Markup Language (SBML): Language Specification for Level 3 Version 2 Core Release 2”. eng. In: *Journal of Integrative Bioinformatics* 16.2 (June 2019), p. 20190021. ISSN: 1613-4516. DOI: 10.1515/jib-2019-0021.

[136] Claudine Chaouiya et al. “SBML qualitative models: a model representation format and infrastructure to foster interactions between qualitative modelling formalisms and tools”. In: *BMC Systems Biology* 7.1 (Dec. 2013), p. 135. ISSN: 1752-0509. DOI: 10.1186/1752-0509-7-135. URL: https://doi.org/10.1186/1752-0509-7-135 (visited on 01/02/2023).

[137] F. Büchel. “Path2Models: large-scale generation of computational models from biochemical pathway maps”. en. In: *BMC Syst. Biol* 7 (2013), p. 116.

[138] S.S. Aghamiri and F. Delaplace. “TaBooN Boolean Network Synthesis Based on Tabu Search”. en. In: *IEEE/ACM Trans. Comput. Biol. Bioinform* (2021), pp. 1–1.

[139] C. Terfve. “CellNOptR: a flexible toolkit to train protein signaling networks to data using multiple logic formalisms”. en. In: *BMC Syst. Biol* 6 (2012), p. 133.

[140] D.-H. Le and Y.-K. Kwon. “NetDS: a Cytoscape plugin to analyze the robustness of dynamics and feedforward/feedback loop structures of biological networks”. en. In: *Bioinformatics* 27 (2011), pp. 2767–2768.

[141] C. Su and J. Pang. “CABEAN: a software for the control of asynchronous Boolean networks”. en. In: *Bioinformatics* 37 (2021), pp. 879–881.

[142] T. Helikar. “The Cell Collective: Toward an open and collaborative approach to systems biology”. en. In: *BMC Syst. Biol* 6 (2012), p. 96.

[143] A. Naldi. “Logical Modeling and Analysis of Cellular Regulatory Networks With GINsim 3.0”. en. In: *Front. Physiol* 9 (2018).

[144] F. Hinkelmann. “ADAM: Analysis of Discrete Models of Biological Systems Using Computer Algebra”. pt. In: *BMC Bioinformatics* 12 (2011), p. 295.

[145] G. Stoll. “MaBoSS 2.0: an environment for stochastic Boolean modeling”. pt. In: *Bioinforma. Oxf. Engl* 33 (2017), pp. 2226–2228.

[146] R.B. Correia. “CANA: A Python Package for Quantifying Control and Canalization in Boolean Networks”. en. In: *Front. Physiol* 9 (2018).

[147] A. Mizera. “ASSA-PBN: An Approximate Steady-State Analyser of Probabilistic Boolean Networks”. en. In: *Automated Technology for Verification and Analysis, Lecture Notes in Computer Science*. Ed. by Finkbeiner In and B. Cham: Springer International Publishing, 2015, pp. 214–220.

[148] D. Benque. “Bma: Visual Tool for Modeling and Analyzing Biological Networks”. en. In: *Computer Aided Verification, Lecture Notes in Computer Science*. Ed. by Madhusudan In, P., and S.A. Seshia. Berlin Heidelberg, Berlin, Heidelberg: Springer, 2012, pp. 686–692.

[149] S.H. Elmeligy Abdelhamid. “GDSCalc: A Web-Based Application for Evaluating Discrete Graph Dynamical Systems”. en. In: *PloS One* 10 (2015), p. 0133660.

[150] A. Paroni. “CABeRNET: a Cytoscape app for augmented Boolean models of gene regulatory NETworks”. en. In: *BMC Bioinformatics* 17 (2016).

[151] A. Bruner and R. Sharan. “A Robustness Analysis of Dynamic Boolean Models of Cellular Circuits”. en. In: *J. Comput. Biol* 27 (2019), pp. 133–143.

[152] J. Béal. “Personalization of Logical Models With Multi-Omics Data Allows Clinical Stratification of Patients”. en. In: *Front. Physiol* 9 (2019).

[153] L. Siegle. “A Boolean network of the crosstalk between IGF and Wnt signaling in aging satellite cells”. en. In: *PloS One* 13 (2018), p. 0195126.

[154] E. Gjerga. *Converting networks to predictive logic models from perturbation signalling data with CellNOpt Bioinformatics*. en. 2020.

[155] J. Saez-Rodriguez. “A Logical Model Provides Insights into T Cell Receptor Signaling”. en. In: *PLOS Comput. Biol* 3 (2007), p. 163.

[156] A. Naldi. “Diversity and Plasticity of Th Cell Types Predicted from Regulatory Network Modelling”. en. In: *PLOS Comput. Biol* 6 (2010), p. 1000912.

[157] B.L. Puniya. “A Mechanistic Computational Model Reveals That Plasticity of CD4+ T Cell Differentiation Is a Function of Cytokine Composition and Dosage”. en. In: *Front. Physiol* 9 (2018).

[158] X.O. Yang. “Molecular antagonism and plasticity of regulatory and inflammatory T cell programs”. en. In: *Immunity* 29 (2008), pp. 44–56.

[159] A. Kimura and T. Kishimoto. “IL-6: regulator of Treg/Th17 balance”. en. In: *Eur. J. Immunol* 40 (2010), pp. 1830–1835.

[160] M. Peine. “Stable T-bet+GATA-3+ Th1/Th2 Hybrid Cells Arise In Vivo, Can Develop Directly from Naive Precursors, and Limit Immunopathologic Inflammation”. en. In: *PLoS Biol* 11 (2013).

[161] Quentin Miagoux et al. “Inference of an Integrative, Executable Network for Rheumatoid Arthritis Combining Data-Driven Machine Learning Approaches and a State-of-the-Art Mechanistic Disease Map”. en. In: *Journal of Personalized Medicine* 11.8 (Aug. 2021). Number: 8 Publisher: Multidisciplinary Digital Publishing Institute, p. 785. ISSN: 2075-4426. DOI: 10.3390/jpm11080785. URL: https://www.mdpi.com/2075-4426/11/8/785 (visited on 11/27/2022).

[162] J. Luo. “Principles of cancer therapy: oncogene and non-oncogene addiction”. en. In: *Cell* 136 (2009), pp. 823–837.

[163] F. Eduati. “Patient-specific logic models of signaling pathways from screenings on cancer biopsies to prioritize personalized combination therapies”. en. In: *Mol. Syst. Biol* 16 (2020), p. 8664.

[164] I. Irurzun-Arana. “Advanced Boolean modeling of biological networks applied to systems pharmacology”. en. In: *Bioinformatics* 33 (2017), pp. 1040–1048.

[165] Jonas Béal et al. “Personalization of Logical Models With Multi-Omics Data Allows Clinical Stratification of Patients”. In: *Frontiers in Physiology* 9 (2019). ISSN: 1664-042X. URL: https://www.frontiersin.org/articles/10.3389/fphys.2018.01965 (visited on 11/27/2022).

[166] H.F. Fumiã and M.L. Martins. “Boolean Network Model for Cancer Pathways: Predicting Carcinogenesis and Targeted Therapy Outcomes”. en. In: *PLOS ONE* 8 (2013), p. 69008.

[167] M. Z and L. H. “Boolean network-based analysis of the apoptosis network: irreversible apoptosis and stable surviving”. en. In: *J. Theor. Biol* 259 (2009), pp. 760–769.

[168] Curtis C et al. “The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups.” English. In: *Nature* 486.7403 (Apr. 2012), pp. 346–352. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/nature10983. URL: http://europepmc.org/article/MED/22522925 (visited on 02/07/2022).

[169] Gautier Stoll et al. “MaBoSS 2.0: an environment for stochastic Boolean modeling”. In: *Bioinformatics* 33.14 (July 2017), pp. 2226–2228. ISSN: 1367-4803. DOI: 10.1093/bioinformatics/btx123. URL: https://doi.org/10.1093/bioinformatics/btx123 (visited on 11/27/2022).

[170] J. Saez-Rodriguez and N. Blüthgen. “Personalized signaling models for personalized treatments”. en. In: *Mol. Syst. Biol* 16 (2020), p. 9042.

[171] B. Klinger. “Network quantification of EGFR signaling unveils potential for targeted combination therapy”. en. In: *Mol. Syst. Biol* 9 (2013), p. 673.

[172] Arnau Montagud et al. “Patient-specific Boolean models of signalling networks guide personalised treatments”. In: *eLife* 11 (Feb. 2022). Ed. by Jennifer Flegg, Aleksandra M Walczak, and Réka Albert. Publisher: eLife Sciences Publications, Ltd, e72626. ISSN: 2050-084X. DOI: 10.7554/eLife.72626. URL: https://doi.org/10.7554/eLife.72626 (visited on 11/27/2022).

[173] J. Saez-Rodriguez. “A Logical Model Provides Insights into T Cell Receptor Signaling”. en. In: *PLOS Comput. Biol* 3 (2007), p. 163.

[174] B.L. Puniya. “A Mechanistic Computational Model Reveals That Plasticity of CD4+ T Cell Differentiation Is a Function of Cytokine Composition and Dosage”. en. In: *Front. Physiol* 9 (2018).

[175] Å. Flobak. “Discovery of Drug Synergies in Gastric Cancer Cells Predicted by Logical Modeling”. en. In: *PLOS Comput. Biol* 11 (2015), p. 1004426.

[176] J. Luo. “Principles of cancer therapy: oncogene and non-oncogene addiction”. en. In: *Cell* 136 (2009), pp. 823–837.

[177] Zhongxing Mai and Haiyan Liu. “Boolean network-based analysis of the apoptosis network: irreversible apoptosis and stable surviving”. eng. In: *Journal of Theoretical Biology* 259.4 (Aug. 2009), pp. 760–769. ISSN: 1095-8541. DOI: 10.1016/j.jtbi.2009.04.024.

[178] H. Sizek. “Boolean model of growth signaling, cell cycle and apoptosis predicts the molecular mechanism of aberrant cell cycle progression driven by hyperactive PI3K”. en. In: *PLOS Comput. Biol* 15 (2019), p. 1006402.

[179] Julian P. T. Higgins et al. “The Cochrane Collaboration’s tool for assessing risk of bias in randomised trials”. en. In: *BMJ* 343 (Oct. 2011). Publisher: British Medical Journal Publishing Group Section: Research Methods & Reporting, p. d5928. ISSN: 0959-8138, 1468-5833. DOI: 10.1136/bmj.d5928. URL: https://www.bmj.com/content/343/bmj.d5928 (visited on 01/07/2023).

[180] Carson Ka-Lok Lo, Dominik Mertz, and Mark Loeb. “Newcastle-Ottawa Scale: comparing reviewers’ to authors’ assessments”. In: *BMC Medical Research Methodology* 14.1 (Apr. 2014), p. 45. ISSN: 1471-2288. DOI: 10.1186/1471-2288-14-45. URL: https://doi.org/10.1186/1471-2288-14-45 (visited on 01/07/2023).

[181] Avi Ma’ayan. “Introduction to Network Analysis in Systems Biology”. In: *Science signaling* 4.190 (Sept. 2011), tr5. ISSN: 1937-9145. DOI: 10.1126/scisignal.2001965. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3196357/ (visited on 12/18/2022).

[182] Minoo Ashtiani et al. “A systematic survey of centrality measures for protein-protein interaction networks”. In: *BMC Systems Biology* 12.1 (July 2018), p. 80. ISSN: 1752-0509. DOI: 10.1186/s12918-018-0598-2. URL: https://doi.org/10.1186/s12918-018-0598-2 (visited on 12/18/2022).

[183] Ahmed Hemedan. *Ahmed Hemedan / Boolean modelling of PD-Thesis · GitLab*. en. Jan. 2023. URL: https://git-r3lab.uni.lu/Ahmed7emdan/applications-of-boolean-modelling-to-study-and-stratify-dynamics-of-a-complex-disease (visited on 01/04/2023).

[184] M. Hopfensitz. “Multiscale binarization of gene expression data for reconstructing Boolean networks”. en. In: *IEEE/ACM Trans. Comput. Biol. Bioinform* 9 (2012), pp. 487–498.

[185] K. Klemm and S. Bornholdt. “Stable and unstable attractors in Boolean networks”. en. In: *Phys. Rev. E* 72 (2005), p. 055101.

[186] Abhishek Garg et al. “Synchronous versus asynchronous modeling of gene regulatory networks”. eng. In: *Bioinformatics (Oxford, England)* 24.17 (Sept. 2008), pp. 1917–1925. ISSN: 1367-4811. DOI: 10.1093/bioinformatics/btn336.

[187] Amanda Ross and Victor L. Willson. “Paired Samples T-Test”. en. In: *Basic and Advanced Statistical Tests: Writing Results Sections and Creating Tables and Figures*. Ed. by Amanda Ross and Victor L. Willson. Rotterdam: SensePublishers, 2017, pp. 17–19. ISBN: 978-94-6351-086-8. DOI: 10.1007/978-94-6351-086-8_4. URL: https://doi.org/10.1007/978-94-6351-086-8_4 (visited on 01/07/2023).

[188] Denise Rey and Markus Neuhäuser. “Wilcoxon-Signed-Rank Test”. en. In: *International Encyclopedia of Statistical Science*. Ed. by Miodrag Lovric. Berlin, Heidelberg: Springer, 2011, pp. 1658–1659. ISBN: 978-3-642-04898-2. DOI: 10.1007/978-3-642-04898-2_616. URL: https://doi.org/10.1007/978-3-642-04898-2_616 (visited on 01/07/2023).

[189] “Pearson’s Correlation Coefficient”. en. In: *Encyclopedia of Public Health*. Ed. by Wilhelm Kirch. Dordrecht: Springer Netherlands, 2008, pp. 1090–1091. ISBN: 978-1-4020-5614-7. DOI: 10.1007/978-1-4020-5614-7_2569. URL: https://doi.org/10.1007/978-1-4020-5614-7_2569 (visited on 12/31/2022).

[190] Jacob Cohen. *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed. New York: Routledge, July 1988. ISBN: 978-0-203-77158-7. DOI: 10.4324/9780203771587.

[191] John Ruscio. “A probability-based measure of effect size: Robustness to base rates and other factors.” en. In: *Psychological Methods* 13.1 (2008), pp. 19–30. ISSN: 1939-1463, 1082-989X. DOI: 10.1037/1082-989X.13.1.19. URL: http://doi.apa.org/getdoi.cfm?doi=10.1037/1082-989X.13.1.19 (visited on 11/27/2022).

[192] Kenneth O. McGraw and S. P. Wong. “A common language effect size statistic”. In: *Psychological Bulletin* 111 (1992). Place: US Publisher: American Psychological Association, pp. 361–365. ISSN: 1939-1455. DOI: 10.1037/0033-2909.111.2.361.

[193] Hsi-Yuan Huang et al. “miRTarBase update 2022: an informative resource for experimentally validated miRNA–target interactions”. In: *Nucleic Acids Research* 50.D1 (Jan. 2022), pp. D222–D230. ISSN: 0305-1048. DOI: 10.1093/nar/gkab1079. URL: https://doi.org/10.1093/nar/gkab1079 (visited on 11/29/2022).

[194] Kazuhiro A. Fujita et al. “Integrating Pathways of Parkinson’s Disease in a Molecular Interaction Map”. en. In: *Molecular Neurobiology* 49.1 (Feb. 2014), pp. 88–102. ISSN: 1559-1182. DOI: 10.1007/s12035-013-8489-4. URL: https://doi.org/10.1007/s12035-013-8489-4 (visited on 01/03/2023).

[195] Gioacchino P. Marceca et al. “MiREDiBase, a manually curated database of validated and putative editing events in microRNAs”. en. In: *Scientific Data* 8.1 (Aug. 2021). Number: 1 Publisher: Nature Publishing Group, p. 199. ISSN: 2052-4463. DOI: 10.1038/s41597-021-00979-8. URL: https://www.nature.com/articles/s41597-021-00979-8 (visited on 11/27/2022).

[196] Eduardo Andrés-León et al. “miRGate: a curated database of human, mouse and rat miRNA–mRNA targets”. In: *Database* 2015 (Jan. 2015), bav035. ISSN: 1758-0463. DOI: 10.1093/database/bav035. URL: https://doi.org/10.1093/database/bav035 (visited on 11/27/2022).

[197] Zhou Huang et al. “HMDD v3.0: a database for experimentally supported human microRNA-disease associations”. eng. In: *Nucleic Acids Research* 47.D1 (Jan. 2019), pp. D1013–D1017. ISSN: 1362-4962. DOI: 10.1093/nar/gky1010.

[198] Xuping Jiang et al. “Screening of Parkinson’s Differential MicroRNA Based on GEO Database and Its Clinical Verification”. eng. In: *BioMed Research International* 2021 (2021), p. 8171236. ISSN: 2314-6141. DOI: 10.1155/2021/8171236.

[199] Enrico Glaab et al. “EnrichNet: network-based gene set enrichment analysis”. In: *Bioinformatics* 28.18 (Sept. 2012), pp. i451–i457. ISSN: 1367-4803. DOI: 10.1093/bioinformatics/bts389. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3436816/ (visited on 01/04/2023).

[200] Antonio Fabregat et al. “The Reactome Pathway Knowledgebase”. In: *Nucleic Acids Research* 46.Database issue (Jan. 2018), pp. D649–D655. ISSN: 0305-1048. DOI: 10.1093/nar/gkx1132. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5753187/ (visited on 01/04/2023).

[201] David J Lynn et al. “InnateDB: facilitating systems-level analyses of the mammalian innate immune response”. In: *Molecular Systems Biology* 4 (Sept. 2008), p. 218. ISSN: 1744-4292. DOI: 10.1038/msb.2008.55. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2564732/ (visited on 01/04/2023).

[202] THE GTEX CONSORTIUM. “The GTEx Consortium atlas of genetic regulatory effects across human tissues”. In: *Science* 369.6509 (Sept. 2020). Publisher: American Association for the Advancement of Science, pp. 1318–1330. DOI: 10.1126/science.aaz1776. URL: https://www.science.org/doi/10.1126/science.aaz1776 (visited on 01/04/2023).

[203] Peter J. Thul and Cecilia Lindskog. “The human protein atlas: A spatial map of the human proteome”. eng. In: *Protein Science: A Publication of the Protein Society* 27.1 (Jan. 2018), pp. 233–244. ISSN: 1469-896X. DOI: 10.1002/pro.3307.

[204] Jonas Kristoffer Lindeløv. *mcp: An R Package for Regression With Multiple Change Points*. en-us. Jan. 2020. DOI: 10.31219/osf.io/fzqxv. URL: https://osf.io/fzqxv/ (visited on 11/27/2022).

[205] Toni Giorgino. “Computing and Visualizing Dynamic Time Warping Alignments in *R* : The **dtw** Package”. en. In: *Journal of Statistical Software* 31.7 (2009). ISSN: 1548-7660. DOI: 10.18637/jss.v031.i07. URL: http://www.jstatsoft.org/v31/i07/ (visited on 11/27/2022).

[206] Yanhua Shen et al. “Regulation of SIRT3 on mitochondrial functions and oxidative stress in Parkinson’s disease”. en. In: *Biomedicine & Pharmacotherapy* 132 (Dec. 2020), p. 110928. ISSN: 0753-3322. DOI: 10.1016/j.biopha.2020.110928. URL: https://www.sciencedirect.com/science/article/pii/S0753332220311203 (visited on 11/03/2021).

[207] Eunyong Ahn et al. “Temporal fluxomics reveals oscillations in TCA cycle flux throughout the mammalian cell cycle”. In: *Molecular Systems Biology* 13.11 (Nov. 2017), p. 953. ISSN: 1744-4292. DOI: 10.15252/msb.20177763. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5731346/ (visited on 11/21/2022).

[208] Peter W. Stacpoole. “The pyruvate dehydrogenase complex as a therapeutic target for age-related diseases”. eng. In: *Aging Cell* 11.3 (June 2012), pp. 371–377. ISSN: 1474-9726. DOI: 10.1111/j.1474-9726.2012.00805.x.

[209] Consolato Sergi and Bindu Parayil Sankaran. “Succinic Semialdehyde Dehydrogenase Deficiency”. eng. In: *StatPearls*. Treasure Island (FL): StatPearls Publishing, 2021. URL: http://www.ncbi.nlm.nih.gov/books/NBK560724/ (visited on 11/02/2021).

[210] Qingli Shi et al. “Inactivation and reactivation of the mitochondrial α-ketoglutarate dehydrogenase complex”. eng. In: *The Journal of Biological Chemistry* 286.20 (May 2011), pp. 17640–17648. ISSN: 1083-351X. DOI: 10.1074/jbc.M110.203018.

[211] G. E. Gibson et al. “Deficits in a tricarboxylic acid cycle enzyme in brains from patients with Parkinson’s disease”. eng. In: *Neurochemistry International* 43.2 (July 2003), pp. 129–135. ISSN: 0197-0186. DOI: 10.1016/s0197-0186(02)00225-5.

[212] Hyunjin Kim et al. “IDH2 deficiency promotes mitochondrial dysfunction and dopaminergic neurotoxicity: implications for Parkinson’s disease”. eng. In: *Free Radical Research* 50.8 (Aug. 2016), pp. 853–860. ISSN: 1029-2470. DOI: 10.1080/10715762.2016.1185519.

[213] Bruna Barneda-Zahonero et al. “Nurr1 protein is required for N-methyl-D-aspartic acid (NMDA) receptor-mediated neuronal survival”. eng. In: *The Journal of Biological Chemistry* 287.14 (Mar. 2012), pp. 11351–11362. ISSN: 1083-351X. DOI: 10.1074/jbc.M111.272427.

[214] Colleen A. Stoyas et al. “Nicotinamide Pathway-Dependent Sirt1 Activation Restores Calcium Homeostasis to Achieve Neuroprotection in Spinocerebellar Ataxia Type 7”. eng. In: *Neuron* 105.4 (Feb. 2020), 630–644.e9. ISSN: 1097-4199. DOI: 10.1016/j.neuron.2019.11.019.

[215] Xuan Li et al. “The Critical Role of SIRT1 in Parkinson’s Disease: Mechanism and Therapeutic Considerations”. In: *Aging and Disease* 11.6 (Dec. 2020), pp. 1608–1622. ISSN: 2152-5250. DOI: 10.14336/AD.2020.0216. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7673849/ (visited on 01/05/2023).

[216] J. Jankovic, S. Chen, and W. D. Le. "The role of Nurr1 in the development of dopaminergic neurons and Parkinson's disease". eng. In: *Progress in Neurobiology* 77.1-2 (Oct. 2005), pp. 128–138. ISSN: 0301-0082. DOI: 10.1016/j.pneurobio.2005.09.001.

[217] Banafsheh Kadkhodaei et al. "Transcription factor Nurr1 maintains fiber integrity and nuclear-encoded mitochondrial gene expression in dopamine neurons". eng. In: *Proceedings of the National Academy of Sciences of the United States of America* 110.6 (Feb. 2013), pp. 2360–2365. ISSN: 1091-6490. DOI: 10.1073/pnas.1221077110.

[218] Frank M. J. Jacobs et al. "Pitx3 potentiates Nurr1 in dopamine neuron terminal differentiation through release of SMRT-mediated repression". eng. In: *Development (Cambridge, England)* 136.4 (Feb. 2009), pp. 531–540. ISSN: 0950-1991. DOI: 10.1242/dev.029769.

[219] Marta Madureira, Natalie Connor-Robson, and Richard Wade-Martins. ""LRRK2: Autophagy and Lysosomal Activity"". eng. In: *Frontiers in Neuroscience* 14 (2020), p. 498. ISSN: 1662-4548. DOI: 10.3389/fnins.2020.00498.

[220] José M. Bravo-San Pedro et al. "The LRRK2 G2019S mutant exacerbates basal autophagy through activation of the MEK/ERK pathway". eng. In: *Cellular and molecular life sciences: CMLS* 70.1 (Jan. 2013), pp. 121–136. ISSN: 1420-9071. DOI: 10.1007/s00018-012-1061-y.

[221] Yoana Rabanal-Ruiz, Elsje G. Otten, and Viktor I. Korolchuk. "mTORC1 as the main gateway to autophagy". eng. In: *Essays in Biochemistry* 61.6 (Dec. 2017), pp. 565–584. ISSN: 1744-1358. DOI: 10.1042/EBC20170027.

[222] Matthew P. McCabe et al. "Genetic inactivation of mTORC1 or mTORC2 in neurons reveals distinct functions in glutamatergic synaptic transmission". eng. In: *eLife* 9 (Mar. 2020), e51440. ISSN: 2050-084X. DOI: 10.7554/eLife.51440.

[223] Zhou Zhu et al. "Balancing mTOR Signaling and Autophagy in the Treatment of Parkinson's Disease". eng. In: *International Journal of Molecular Sciences* 20.3 (Feb. 2019), E728. ISSN: 1422-0067. DOI: 10.3390/ijms20030728.

[224] Berenice Franco-Juárez et al. "TFEB; Beyond Its Role as an Autophagy and Lysosomes Regulator". en. In: *Cells* 11.19 (Jan. 2022). Number: 19 Publisher: Multidisciplinary Digital Publishing Institute, p. 3153. ISSN: 2073-4409. DOI: 10.3390/cells11193153. URL: https://www.mdpi.com/2073-4409/11/19/3153 (visited on 12/26/2022).

[225] Mickael Decressac et al. "TFEB-mediated autophagy rescues midbrain dopamine neurons from α-synuclein toxicity". In: *Proceedings of the National Academy of Sciences* 110.19 (May 2013). Publisher: Proceedings of the National Academy of Sciences, E1817–E1826. DOI: 10.1073/pnas.1305623110. URL: https://www.pnas.org/doi/10.1073/pnas.1305623110 (visited on 12/26/2022).

[226] Tim E. Moors et al. “Therapeutic potential of autophagy-enhancing agents in Parkinson’s disease”. eng. In: *Molecular Neurodegeneration* 12.1 (Jan. 2017), p. 11. ISSN: 1750-1326. DOI: 10.1186/s13024-017-0154-3.

[227] Yongle Li, Jun Zhang, and Chunxiang Yang. “UNC-51-like kinase 1 blocks S6k1 phosphorylation contributes to neurodegeneration in Parkinson’s disease model in vitro”. eng. In: *Biochemical and Biophysical Research Communications* 459.2 (Apr. 2015), pp. 196–200. ISSN: 1090-2104. DOI: 10.1016/j.bbrc.2015.02.008.

[228] Ernest Palomer, Johanna Buechler, and Patricia C. Salinas. “Wnt Signaling Deregulation in the Aging and Alzheimer’s Brain”. In: *Frontiers in Cellular Neuroscience* 13 (2019), p. 227. ISSN: 1662-5102. DOI: 10.3389/fncel.2019.00227. URL: https://www.frontiersin.org/article/10.3389/fncel.2019.00227 (visited on 11/04/2021).

[229] Miguel A. Hermida, J. Dinesh Kumar, and Nick R. Leslie. “GSK3 and its interactions with the PI3K/AKT/mTOR signalling network”. en. In: *Advances in Biological Regulation*. Diverse Roles of Glycogen Synthase Kinase 3 (GSK-3) In Signal Transduction and Human Health 65 (Aug. 2017), pp. 5–15. ISSN: 2212-4926. DOI: 10.1016/j.jbior.2017.06.003. URL: https://www.sciencedirect.com/science/article/pii/S2212492617301240 (visited on 12/26/2022).

[230] Emilda Pino et al. “FOXO3 determines the accumulation of α-synuclein and controls the fate of dopaminergic neurons in the substantia nigra”. eng. In: *Human Molecular Genetics* 23.6 (Mar. 2014), pp. 1435–1452. ISSN: 1460-2083. DOI: 10.1093/hmg/ddt530.

[231] Zhiyong Cheng. “FoxO transcription factors in mitochondrial homeostasis”. eng. In: *The Biochemical Journal* 479.4 (Feb. 2022), pp. 525–536. ISSN: 1470-8728. DOI: 10.1042/BCJ20210777.

[232] Candida Fasano et al. “FOXO3a from the Nucleus to the Mitochondria: A Round Trip in Cellular Stress Response”. eng. In: *Cells* 8.9 (Sept. 2019), E1110. ISSN: 2073-4409. DOI: 10.3390/cells8091110.

[233] Xu Hou et al. “Autophagy in Parkinson’s Disease”. en. In: *Journal of Molecular Biology*. Autophagy in Neurodegenerative Diseases 432.8 (Apr. 2020), pp. 2651–2672. ISSN: 0022-2836. DOI: 10.1016/j.jmb.2020.01.037. URL: https://www.sciencedirect.com/science/article/pii/S0022283620301030 (visited on 12/26/2022).

[234] Federica Albanese, Salvatore Novello, and Michele Morari. “Autophagy and LRRK2 in the Aging Brain”. In: *Frontiers in Neuroscience* 13 (2019). ISSN: 1662-453X. URL: https://www.frontiersin.org/articles/10.3389/fnins.2019.01352 (visited on 01/07/2023).

[235] C. Alexander Boecker et al. “Increased LRRK2 kinase activity alters neuronal autophagy by disrupting the axonal transport of autophagosomes”. en. In: *Current Biology* 31.10 (May 2021), 2140–2154.e6. ISSN: 0960-9822. DOI: 10.1016/j.cub.2021.02.061. URL: https://www.sciencedirect.com/science/article/pii/S0960982221003055 (visited on 01/07/2023).

[236] Lifen Zhang et al. “Age-related changes in dopamine signaling in Nurr1 deficient mice as a model of Parkinson’s disease”. eng. In: *Neurobiology of Aging* 33.5 (May 2012), 1001.e7–16. ISSN: 1558-1497. DOI: 10.1016/j.neurobiolaging.2011.03.022.

[237] Xu-Xu Zhuang et al. “Pharmacological enhancement of TFEB-mediated autophagy alleviated neuronal death in oxidative stress-induced Parkinson’s disease models”. en. In: *Cell Death & Disease* 11.2 (Feb. 2020). Number: 2 Publisher: Nature Publishing Group, pp. 1–18. ISSN: 2041-4889. DOI: 10.1038/s41419-020-2322-6. URL: https://www.nature.com/articles/s41419-020-2322-6 (visited on 12/26/2022).

[238] Bruno Jorge de Andrade Silva et al. “Autophagy Is an Innate Mechanism Associated with Leprosy Polarization”. In: *PLoS Pathogens* 13.1 (Jan. 2017), e1006103. ISSN: 1553-7366. DOI: 10.1371/journal.ppat.1006103. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5215777/ (visited on 01/05/2023).

[239] Klodjan Stafa et al. “GTPase activity and neuronal toxicity of Parkinson’s disease-associated LRRK2 is regulated by ArfGAP1”. eng. In: *PLoS genetics* 8.2 (2012), e1002526. ISSN: 1553-7404. DOI: 10.1371/journal.pgen.1002526.

[240] Anelya Kh Alieva et al. “miRNA expression is highly sensitive to a drug therapy in Parkinson’s disease”. eng. In: *Parkinsonism & Related Disorders* 21.1 (Jan. 2015), pp. 72–74. ISSN: 1873-5126. DOI: 10.1016/j.parkreldis.2014.10.018.

[241] Teresa Botta-Orfila et al. “Identification of blood serum micro-RNAs associated with idiopathic and LRRK2 Parkinson’s disease”. eng. In: *Journal of Neuroscience Research* 92.8 (Aug. 2014), pp. 1071–1077. ISSN: 1097-4547. DOI: 10.1002/jnr.23377.

[242] Lucía F. Cardo et al. “Profile of microRNAs in the plasma of Parkinson’s disease patients and healthy controls”. eng. In: *Journal of Neurology* 260.5 (May 2013), pp. 1420–1422. ISSN: 1432-1459. DOI: 10.1007/s00415-013-6900-8.

[243] Jieshan Chi et al. “Integrated Analysis and Identification of Novel Biomarkers in Parkinson’s Disease”. eng. In: *Frontiers in Aging Neuroscience* 10 (2018), p. 178. ISSN: 1663-4365. DOI: 10.3389/fnagi.2018.00178.

[244] Haixia Ding et al. “Identification of a panel of five serum miRNAs as a biomarker for Parkinson’s disease”. eng. In: *Parkinsonism & Related Disorders* 22 (Jan. 2016), pp. 68–73. ISSN: 1873-5126. DOI: 10.1016/j.parkreldis.2015.11.014.

[245] Andrew G. Hoss et al. “microRNA Profiles in Parkinson’s Disease Prefrontal Cortex”. In: *Frontiers in Aging Neuroscience* 8 (Mar. 2016), p. 36. ISSN: 1663-4365. DOI: 10.3389/fnagi.2016.00036. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4772525/ (visited on 11/12/2022).

[246] Sok Kean Khoo et al. “Plasma-based circulating MicroRNA biomarkers for Parkinson’s disease”. eng. In: *Journal of Parkinson’s Disease* 2.4 (2012), pp. 321–331. ISSN: 1877-718X. DOI: 10.3233/JPD-012144.

[247] Madalena Martins et al. “Convergence of miRNA expression profiling, α-synuclein interacton and GWAS in Parkinson’s disease”. eng. In: *PloS One* 6.10 (2011), e25443. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0025443.

[248] Regina Margis, Rogério Margis, and Carlos R. M. Rieder. “Identification of blood microRNAs associated to Parkinsonĭs disease”. eng. In: *Journal of Biotechnology* 152.3 (Mar. 2011), pp. 96–101. ISSN: 1873-4863. DOI: 10.1016/j.jbiotec.2011.01.023.

[249] Alice Serafin et al. “Overexpression of blood microRNAs 103a, 30b, and 29a in L-dopa-treated patients with PD”. eng. In: *Neurology* 84.7 (Feb. 2015), pp. 645–653. ISSN: 1526-632X. DOI: 10.1212/WNL.0000000000001258.

[250] Lilach Soreq et al. “Small RNA sequencing-microarray analyses in Parkinson leukocytes reveal deep brain stimulation-induced splicing changes that classify brain region transcriptomes”. In: *Frontiers in Molecular Neuroscience* 6 (2013). ISSN: 1662-5099. URL: https://www.frontiersin.org/articles/10.3389/fnmol.2013.00010 (visited on 11/12/2022).

[251] Stylianos Ravanidis et al. “Validation of differentially expressed brain-enriched microRNAs in the plasma of PD patients”. In: *Annals of Clinical and Translational Neurology* 7.9 (Aug. 2020), pp. 1594–1607. ISSN: 2328-9503. DOI: 10.1002/acn3.51146. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7480914/ (visited on 11/12/2022).

[252] Annamaria Vallelunga et al. “Serum miR-96-5P and miR-339-5P Are Potential Biomarkers for Multiple System Atrophy and Parkinson’s Disease”. In: *Frontiers in Aging Neuroscience* 13 (2021). ISSN: 1663-4365. URL: https://www.frontiersin.org/articles/10.3389/fnagi.2021.632891 (visited on 11/12/2022).

[253] Annamaria Vallelunga et al. “Identification of circulating microRNAs for the differential diagnosis of Parkinson’s disease and Multiple System Atrophy”. eng. In: *Frontiers in Cellular Neuroscience* 8 (2014), p. 156. ISSN: 1662-5102. DOI: 10.3389/fncel.2014.00156.

[254] Xiang-Yang Cao et al. “MicroRNA biomarkers of Parkinson’s disease in serum exosome-like microvesicles”. eng. In: *Neuroscience Letters* 644 (Mar. 2017), pp. 94–99. ISSN: 1872-7972. DOI: 10.1016/j.neulet.2017.02.045.

[255] Franciele Cascaes da Silva et al. “microRNAs involved in Parkinson’s disease: A systematic review”. In: *Molecular Medicine Reports* 14.5 (Nov. 2016). Publisher: Spandidos Publications, pp. 4015–4022. ISSN: 1791-2997. DOI: 10.3892/mmr.2016.5759. URL: https://www.spandidos-publications.com/10.3892/mmr.2016.5759 (visited on 11/12/2022).

[256] Arnon Blum et al. “MicroRNA-423 may regulate diabetic vasculopathy”. eng. In: *Clinical and Experimental Medicine* 19.4 (Nov. 2019), pp. 469–477. ISSN: 1591-9528. DOI: 10.1007/s10238-019-00573-8.

[257] Xuegui Zhou, Cuiping Xiang, and Xiaoxia Zheng. “miR-132 serves as a diagnostic biomarker in gestational diabetes mellitus and its regulatory effect on trophoblast cell viability”. eng. In: *Diagnostic Pathology* 14.1 (Oct. 2019), p. 119. ISSN: 1746-1596. DOI: 10.1186/s13000-019-0899-9.

[258] Hassan Mziaut et al. “MiR-132 controls pancreatic beta cell proliferation and survival through Pten/Akt/Foxo3 signaling”. eng. In: *Molecular Metabolism* 31 (Jan. 2020), pp. 150–162. ISSN: 2212-8778. DOI: 10.1016/j.molmet.2019.11.012.

[259] Ghada Al-Kafaji et al. “Peripheral blood microRNA-15a is a potential biomarker for type 2 diabetes mellitus and pre-diabetes”. eng. In: *Molecular Medicine Reports* 12.5 (Nov. 2015), pp. 7485–7490. ISSN: 1791-3004. DOI: 10.3892/mmr.2015.4416.

[260] Azadeh Houshmand-Oeregaard et al. “Increased expression of microRNA-15a and microRNA-15b in skeletal muscle from adult offspring of women with diabetes in pregnancy”. eng. In: *Human Molecular Genetics* 27.10 (May 2018), pp. 1763–1771. ISSN: 1460-2083. DOI: 10.1093/hmg/ddy085.

[261] Xiaochen Bai et al. “Downregulation of blood serum microRNA 29 family in patients with Parkinson’s disease”. eng. In: *Scientific Reports* 7.1 (July 2017), p. 5411. ISSN: 2045-2322. DOI: 10.1038/s41598-017-03887-3.

[262] Julie Massart et al. “Altered miR-29 Expression in Type 2 Diabetes Influences Glucose and Lipid Metabolism in Skeletal Muscle”. eng. In: *Diabetes* 66.7 (July 2017), pp. 1807–1818. ISSN: 1939-327X. DOI: 10.2337/db17-0141.

[263] Suh Yee Goh et al. “Role of MicroRNAs in Parkinson’s Disease”. eng. In: *International Journal of Molecular Sciences* 20.22 (Nov. 2019), E5649. ISSN: 1422-0067. DOI: 10.3390/ijms20225649.

[264] James Dooley et al. “The microRNA-29 Family Dictates the Balance Between Homeostatic and Pathological Glucose Handling in Diabetes and Obesity”. eng. In: *Diabetes* 65.1 (Jan. 2016), pp. 53–61. ISSN: 1939-327X. DOI: 10.2337/db15-0770.

[265] Paulami Chatterjee et al. “Studying the system-level involvement of microRNAs in Parkinson’s disease”. eng. In: *PloS One* 9.4 (2014), e93751. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0093751.

[266] Dan Ye et al. “Plasma miR-17, miR-20a, miR-20b and miR-122 as potential biomarkers for diagnosis of NAFLD in type 2 diabetes mellitus patients”. eng. In: *Life Sciences* 208 (Sept. 2018), pp. 201–207. ISSN: 1879-0631. DOI: 10.1016/j.lfs.2018.07.029.

[267] Carmen Pheiffer et al. “Decreased Expression of Circulating miR-20a-5p in South African Women with Gestational Diabetes Mellitus”. eng. In: *Molecular Diagnosis & Therapy* 22.3 (June 2018), pp. 345–352. ISSN: 1179-2000. DOI: 10.1007/s40291-018-0325-0.

[268] Cristina Barbagallo et al. “Specific Signatures of Serum miRNAs as Potential Biomarkers to Discriminate Clinically Similar Neurodegenerative and Vascular-Related Diseases”. eng. In: *Cellular and Molecular Neurobiology* 40.4 (May 2020), pp. 531–546. ISSN: 1573-6830. DOI: 10.1007/s10571-019-00751-y.

[269] Rosalba Senese et al. “miR-22-3p is involved in gluconeogenic pathway modulated by 3,5-diiodo-L-thyronine (T2)”. eng. In: *Scientific Reports* 9.1 (Nov. 2019), p. 16645. ISSN: 2045-2322. DOI: 10.1038/s41598-019-53019-2.

[270] Bridget Martinez and Philip V. Peplow. “MicroRNAs in Parkinson’s disease and emerging therapeutic targets”. eng. In: *Neural Regeneration Research* 12.12 (Dec. 2017), pp. 1945–1959. ISSN: 1673-5374. DOI: 10.4103/1673-5374.221147.

[271] Ying-Zhi Liang et al. “Identification of stress-related microRNA biomarkers in type 2 diabetes mellitus: A systematic review and meta-analysis”. eng. In: *Journal of Diabetes* 12.9 (Sept. 2020), pp. 633–644. ISSN: 1753-0407. DOI: 10.1111/1753-0407.12643.

[272] Lin Xihua et al. “Circulating miR-143-3p inhibition protects against insulin resistance in Metabolic Syndrome via targeting of the insulin-like growth factor 2 receptor”. eng. In: *Translational Research: The Journal of Laboratory and Clinical Medicine* 205 (Mar. 2019), pp. 33–43. ISSN: 1878-1810. DOI: 10.1016/j.trsl.2018.09.006.

[273] Yanmei Chen et al. “Increased Salivary microRNAs That Regulate DJ-1 Gene Expression as Potential Markers for Parkinson’s Disease”. eng. In: *Frontiers in Aging Neuroscience* 12 (2020), p. 210. ISSN: 1663-4365. DOI: 10.3389/fnagi.2020.00210.

[274] Chengcheng Cui et al. “miR-145 Regulates Diabetes-Bone Marrow Stromal Cell-Induced Neurorestorative Effects in Diabetes Stroke Rats”. eng. In: *Stem Cells Translational Medicine* 5.12 (Dec. 2016), pp. 1656–1667. ISSN: 2157-6564. DOI: 10.5966/sctm.2015-0349.

[275] Xiong Zhang et al. “Reduced Circulating Levels of miR-433 and miR-133b Are Potential Biomarkers for Parkinson’s Disease”. eng. In: *Frontiers in Cellular Neuroscience* 11 (2017), p. 170. ISSN: 1662-5102. DOI: 10.3389/fncel.2017.00170.

[276] Xie Y et al. “Urinary Exosomal MicroRNA Profiling in Incipient Type 2 Diabetic Kidney Disease”. en. In: *Journal of diabetes research* 2017 (2017). Publisher: J Diabetes Res. ISSN: 2314-6753. DOI: 10.1155/2017/6978984. URL: https://pubmed.ncbi.nlm.nih.gov/29038788/ (visited on 11/22/2022).

[277] Mahsa Rostamian Delavar et al. “Differential expression of miR-34a, miR-141, and miR-9 in MPP+-treated differentiated PC12 cells as a model of Parkinson’s disease”. eng. In: *Gene* 662 (July 2018), pp. 54–65. ISSN: 1879-0038. DOI: 10.1016/j.gene.2018.04.010.

[278] Ioanna Kokkinopoulou et al. “Decreased expression of microRNAs targeting type-2 diabetes susceptibility genes in peripheral blood of patients and predisposed individuals”. eng. In: *Endocrine* 66.2 (Nov. 2019), pp. 226–239. ISSN: 1559-0100. DOI: 10.1007/s12020-019-02062-0.

[279] Anna-Elisa Roser et al. “miR-182-5p and miR-183-5p Act as GDNF Mimics in Dopaminergic Midbrain Neurons”. eng. In: *Molecular Therapy. Nucleic Acids* 11 (June 2018), pp. 9–22. ISSN: 2162-2531. DOI: 10.1016/j.omtn.2018.01.005.

[280] Cecil Jack Weale et al. “Circulating miR-30a-5p and miR-182-5p in Prediabetes and Screen-Detected Diabetes Mellitus”. eng. In: *Diabetes, Metabolic Syndrome and Obesity: Targets and Therapy* 13 (2020), pp. 5037–5047. ISSN: 1178-7007. DOI: 10.2147/DMSO.S286081.

[281] Nina Mononen et al. “Whole blood microRNA levels associate with glycemic status and correlate with target mRNAs in pathways important to type 2 diabetes”. eng. In: *Scientific Reports* 9.1 (June 2019), p. 8887. ISSN: 2045-2322. DOI: 10.1038/s41598-019-43793-4.

[282] Shujun Wan et al. “Increased serum miR-7 is a promising biomarker for type 2 diabetes mellitus and its microvascular complications”. eng. In: *Diabetes Research and Clinical Practice* 130 (Aug. 2017), pp. 171–179. ISSN: 1872-8227. DOI: 10.1016/j.diabres.2017.06.005.

[283] Jianping Wang et al. “Serum miR-195-5p is upregulated in gestational diabetes mellitus”. eng. In: *Journal of Clinical Laboratory Analysis* 34.8 (Aug. 2020), e23325. ISSN: 1098-2825. DOI: 10.1002/jcla.23325.

[284] Rui-Xian Xing et al. “Down regulation of miR-218, miR-124, and miR-144 relates to Parkinson’s disease via activating NF-κB signaling”. eng. In: *The Kaohsiung Journal of Medical Sciences* 36.10 (Oct. 2020), pp. 786–792. ISSN: 2410-8650. DOI: 10.1002/kjm2.12241.

[285] Rui Yao et al. “Glucose-induced microRNA-218 suppresses the proliferation and promotes the apoptosis of human retinal pigment epithelium cells by targeting RUNX2”. eng. In: *Bioscience Reports* 39.12 (Dec. 2019), BSR20192580. ISSN: 1573-4935. DOI: 10.1042/BSR20192580.

[286] Eiichiro Satake et al. “Circulating miRNA Profiles Associated With Hyperglycemia in Patients With Type 1 Diabetes”. eng. In: *Diabetes* 67.5 (May 2018), pp. 1013–1023. ISSN: 1939-327X. DOI: 10.2337/db17-1207.

[287] Yan Fan et al. “LncRNA BDNF-AS promotes autophagy and apoptosis in MPTP-induced Parkinson’s disease via ablating microRNA-125b-5p”. eng. In: *Brain Research Bulletin* 157 (Apr. 2020), pp. 119–127. ISSN: 1873-2747. DOI: 10.1016/j.brainresbull.2020.02.003.

[288] Cheng-Yong Yu, Chun-Yun Yang, and Zhi-Lian Rui. “MicroRNA-125b-5p improves pancreatic β-cell function through inhibiting JNK signaling pathway by targeting DACT1 in mice with type 2 diabetes mellitus”. eng. In: *Life Sciences* 224 (May 2019), pp. 67–75. ISSN: 1879-0631. DOI: 10.1016/j.lfs.2019.01.031.

[289] Srividya Vasu et al. “MicroRNA Signatures as Future Biomarkers for Diagnosis of Diabetes States”. eng. In: *Cells* 8.12 (Nov. 2019), E1533. ISSN: 2073-4409. DOI: 10.3390/cells8121533.

[290] Olfat G. Shaker et al. “Diagnostic and prognostic role of serum miR-20b, miR-17-3p, HOTAIR, and MALAT1 in diabetic retinopathy”. eng. In: *IUBMB life* 71.3 (Mar. 2019), pp. 310–320. ISSN: 1521-6551. DOI: 10.1002/iub.1970.

[291] Yue Dong, Li-Li Han, and Zhong-Xin Xu. “Suppressed microRNA-96 inhibits iNOS expression and dopaminergic neuron apoptosis through inactivating the MAPK signaling pathway by targeting CACNG5 in mice with Parkinson’s disease”. eng. In: *Molecular Medicine (Cambridge, Mass.)* 24.1 (Nov. 2018), p. 61. ISSN: 1528-3658. DOI: 10.1186/s10020-018-0059-9.

[292] Hyo-Jin Jeong et al. “The induction of miR-96 by mitochondrial dysfunction causes impaired glycogen synthesis through translational repression of IRS-1 in SK-Hep1 cells”. eng. In: *Biochemical and Biophysical Research Communications* 434.3 (May 2013), pp. 503–508. ISSN: 1090-2104. DOI: 10.1016/j.bbrc.2013.03.104.

[293] Liang Zhao and Zhiqin Wang. “MicroRNAs: Game Changers in the Regulation of α-Synuclein in Parkinson’s Disease”. eng. In: *Parkinson’s Disease* 2019 (2019), p. 1743183. ISSN: 2090-8083. DOI: 10.1155/2019/1743183.

[294] Aurora Mazzeo et al. “Functional analysis of miR-21-3p, miR-30b-5p and miR-150-5p shuttled by extracellular vesicles from diabetic subjects reveals their association with diabetic retinopathy”. eng. In: *Experimental Eye Research* 184 (July 2019), pp. 56–63. ISSN: 1096-0007. DOI: 10.1016/j.exer.2019.04.015.

[295] Jie Fu et al. “Regulatory roles of the miR-200 family in neurodegenerative diseases”. eng. In: *Biomedicine & Pharmacotherapy = Biomedecine & Pharmacotherapie* 119 (Nov. 2019), p. 109409. ISSN: 1950-6007. DOI: 10.1016/j.biopha.2019.109409.

[296] Taís S. Assmann et al. “MicroRNA expression profile in plasma from type 1 diabetic patients: Case-control study and bioinformatic analysis”. eng. In: *Diabetes Research and Clinical Practice* 141 (July 2018), pp. 35–46. ISSN: 1872-8227. DOI: 10.1016/j.diabres.2018.03.044.

[297] Anna Elisa Roser et al. “Circulating miRNAs as Diagnostic Biomarkers for Parkinson’s Disease”. In: *Frontiers in Neuroscience* 12 (Sept. 2018), p. 625. ISSN: 1662-4548. DOI: 10.3389/fnins.2018.00625. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6135037/ (visited on 12/26/2022).

[298] Qian Zhang et al. “Acarbose reduces blood glucose by activating miR-10a-5p and miR-664 in diabetic rats”. eng. In: *PloS One* 8.11 (2013), e79697. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0079697.

[299] Sun Y et al. “Inhibition of miR-153, an IL-1β-responsive miRNA, prevents beta cell failure and inflammation-associated diabetes”. en. In: *Metabolism: clinical and experimental* 111 (Oct. 2020). Publisher: Metabolism. ISSN: 1532-8600. DOI: 10.1016/j.metabol.2020.154335. URL: https://pubmed.ncbi.nlm.nih.gov/32795559/ (visited on 11/22/2022).

[300] Masoume Akhbari et al. “Expression Level of Circulating Cell Free miR-155 Gene in Serum of Patients with Diabetic Nephropathy”. eng. In: *Clinical Laboratory* 65.8 (Aug. 2019). ISSN: 1433-6510. DOI: 10.7754/Clin.Lab.2019.190209.

[301] Xu-Wen Tang and Qing-Xin Qin. “miR-335-5p induces insulin resistance and pancreatic islet β-cell secretion in gestational diabetes mellitus mice through VASH1-mediated TGF-β signaling pathway”. eng. In: *Journal of Cellular Physiology* 234.5 (May 2019), pp. 6654–6666. ISSN: 1097-4652. DOI: 10.1002/jcp.27406.

[302] Fusong Jiang et al. “miR-26a Attenuated Bone-Specific Insulin Resistance and Bone Quality in Diabetic Mice”. eng. In: *Molecular Therapy. Nucleic Acids* 20 (June 2020), pp. 459–467. ISSN: 2162-2531. DOI: 10.1016/j.omtn.2020.03.010.

[303] Robert J. A. Frost and Eric N. Olson. “Control of glucose homeostasis and insulin sensitivity by the Let-7 family of microRNAs”. eng. In: *Proceedings of the National Academy of Sciences of the United States of America* 108.52 (Dec. 2011), pp. 21075–21080. ISSN: 1091-6490. DOI: 10.1073/pnas.1118922109.

[304] Paola de Candia et al. “A unique plasma microRNA profile defines type 2 diabetes progression”. eng. In: *PloS One* 12.12 (2017), e0188980. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0188980.

[305] Y.-H. Taguchi and Hsiuying Wang. “Exploring MicroRNA Biomarkers for Parkinson’s Disease from mRNA Expression Profiles”. eng. In: *Cells* 7.12 (Dec. 2018), E245. ISSN: 2073-4409. DOI: 10.3390/cells7120245.

[306] Tao Peng et al. “Long noncoding RNA HAGLROS regulates apoptosis and autophagy in Parkinson’s disease via regulating miR-100/ATG10 axis and PI3K/Akt/mTOR pathway activation”. eng. In: *Artificial Cells, Nanomedicine, and Biotechnology* 47.1 (Dec. 2019), pp. 2764–2774. ISSN: 2169-141X. DOI: 10.1080/21691401.2019.1636805.

[307] Taís S. Assmann et al. “MicroRNA expression profiles and type 1 diabetes mellitus: systematic review and bioinformatic analysis”. eng. In: *Endocrine Connections* 6.8 (Nov. 2017), pp. 773–790. ISSN: 2049-3614. DOI: 10.1530/EC-17-0248.

[308] Ashay D. Bhatwadekar et al. “miR-92a Corrects CD34+ Cell Dysfunction in Diabetes by Modulating Core Circadian Genes Involved in Progenitor Differentiation”. eng. In: *Diabetes* 64.12 (Dec. 2015), pp. 4226–4237. ISSN: 1939-327X. DOI: 10.2337/db15-0521.

[309] Elizabeth M. Kline et al. “Genetic and Environmental Factors in Parkinson’s Disease Converge on Immune Function and Inflammation”. en. In: *Movement Disorders* 36.1 (2021). _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/mds.28411, pp. 25–36. ISSN: 1531-8257. DOI: 10.1002/mds.28411. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/mds.28411 (visited on 12/26/2022).

[310] Xiaolin Zhu et al. “Gene Therapy for Neurodegenerative Disease: Clinical Potential and Directions”. In: *Frontiers in Molecular Neuroscience* 14 (2021). ISSN: 1662-5099. URL: https://www.frontiersin.org/articles/10.3389/fnmol.2021.618171 (visited on 12/26/2022).

[311] Hiroaki Kitano. “Systems biology: a brief overview”. eng. In: *Science (New York, N.Y.)* 295.5560 (Mar. 2002), pp. 1662–1664. ISSN: 1095-9203. DOI: 10.1126/science.1069492.

[312] Levi B. Wood, Ashley R. Winslow, and Samantha Dale Strasser. “Systems Biology of Neurodegenerative Diseases”. In: *Integrative biology : quantitative biosciences from nano to macro* 7.7 (July 2015), pp. 758–775. ISSN: 1757-9694. DOI: 10.1039/c5ib00031a. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4587766/ (visited on 12/10/2022).

[313] Rongting Yue and Abhishek Dutta. “Computational systems biology in disease modeling and control, review and perspectives”. en. In: *npj Systems Biology and Applications* 8.1 (Oct. 2022). Number: 1 Publisher: Nature Publishing Group, pp. 1–16. ISSN: 2056-7189. DOI: 10.1038/s41540-022-00247-4. URL: https://www.nature.com/articles/s41540-022-00247-4 (visited on 12/26/2022).

[314] Bertrand De Meulder et al. “A computational framework for complex disease stratification from multiple large-scale datasets”. In: *BMC Systems Biology* 12.1 (May 2018), p. 60. ISSN: 1752-0509. DOI: 10.1186/s12918-018-0556-z. URL: https://doi.org/10.1186/s12918-018-0556-z (visited on 12/26/2022).

[315] Christian Knüpfer and Clemens Beckstein. “Function of dynamic models in systems biology: linking structure to behaviour”. In: *Journal of Biomedical Semantics* 4 (Oct. 2013), p. 24. ISSN: 2041-1480. DOI: 10.1186/2041-1480-4-24. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3853929/ (visited on 12/26/2022).

[316] Muhammad Usman Sadiq et al. “Model-based stratification of progression along the Alzheimer disease continuum highlights the centrality of biomarker synergies”. In: *Alzheimer’s Research & Therapy* 14.1 (Jan. 2022), p. 16. ISSN: 1758-9193. DOI: 10.1186/s13195-021-00941-1. URL: https://doi.org/10.1186/s13195-021-00941-1 (visited on 12/26/2022).

[317] Sarah Lempriere. “Exosomal microRNA is promising biomarker in PD”. en. In: *Nature Reviews Neurology* 18.2 (Feb. 2022). Number: 2 Publisher: Nature Publishing Group, pp. 65–65. ISSN: 1759-4766. DOI: 10.1038/s41582-021-00611-8. URL: https://www.nature.com/articles/s41582-021-00611-8 (visited on 12/26/2022).

[318] B. Zeigler, H. Praehofer, and T. Kim. “Theory of Modeling and Simulation: Integrating Discrete Event and Continuous Complex Dynamic Systems”. In: Jan. 2000. URL: https://www.semanticscholar.org/paper/Theory-of-Modeling-and-Simulation%3A-Integrating-and-Zeigler-Praehofer/92131a3aed0d72ccfe92364eee87e228c9c773e6 (visited on 12/30/2022).

[319] Antigoni Avramouli. “Validation of Modeling and Simulation Methods in Computational Biology”. en. In: *GeNeDis 2018*. Ed. by Panayiotis Vlamos. Advances in Experimental Medicine and Biology. Cham: Springer International Publishing, 2020, pp. 323–330. ISBN: 978-3-030-32622-7. DOI: 10.1007/978-3-030-32622-7_30.

[320] Paul M. A. Antony et al. “The hallmarks of Parkinson’s disease”. eng. In: *The FEBS journal* 280.23 (Dec. 2013), pp. 5981–5993. ISSN: 1742-4658. DOI: 10.1111/febs.12335.

[321] Adina N. MacMahon Copas et al. “The Pathogenesis of Parkinson’s Disease: A Complex Interplay Between Astrocytes, Microglia, and T Lymphocytes?” In: *Frontiers in Neurology* 12 (2021). ISSN: 1664-2295. URL: https://www.frontiersin.org/articles/10.3389/fneur.2021.666737 (visited on 12/26/2022).

[322] Eleni Kafkia et al. “Operation of a TCA cycle subnetwork in the mammalian nucleus”. In: *Science Advances* 8.35 (Aug. 2022). Publisher: American Association for the Advancement of Science, eabq5206. DOI: 10.1126/sciadv.abq5206. URL: https://www.science.org/doi/10.1126/sciadv.abq5206 (visited on 01/01/2023).

[323] Parith Wongkittichote et al. “Tricarboxylic acid cycle enzyme activities in a mouse model of methylmalonic aciduria”. eng. In: *Molecular Genetics and Metabolism* 128.4 (Dec. 2019), pp. 444–451. ISSN: 1096-7206. DOI: 10.1016/j.ymgme.2019.10.007.

[324] Yong Joo Lee et al. “TCA cycle-independent acetate metabolism via the glyoxylate cycle in Saccharomyces cerevisiae”. en. In: *Yeast* 28.2 (2011). _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/yea.1828, pp. 153–166. ISSN: 1097-0061. DOI: 10.1002/yea.1828. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/yea.1828 (visited on 12/25/2022).

[325] Luciano F. Huergo and Ray Dixon. “The Emergence of 2-Oxoglutarate as a Master Regulator Metabolite”. eng. In: *Microbiology and molecular biology reviews: MMBR* 79.4 (Dec. 2015), pp. 419–435. ISSN: 1098-5557. DOI: 10.1128/MMBR.00038-15.

[326] Liqun Chen et al. “The nuclear receptor 4A family members: mediators in human disease and autophagy”. In: *Cellular & Molecular Biology Letters* 25.1 (Nov. 2020), p. 48. ISSN: 1689-1392. DOI: 10.1186/s11658-020-00241-w. URL: https://doi.org/10.1186/s11658-020-00241-w (visited on 12/25/2022).

[327] Hui-Zhi Long et al. “PI3K/AKT Signal Pathway: A Target of Natural Products in the Prevention and Treatment of Alzheimer’s Disease and Parkinson’s Disease”. In: *Frontiers in Pharmacology* 12 (2021). ISSN: 1663-9812. URL: https://www.frontiersin.org/articles/10.3389/fphar.2021.648636 (visited on 12/26/2022).

[328] Gonçalo Castelo-Branco, Nina Rawal, and Ernest Arenas. “GSK-3beta inhibition/beta-catenin stabilization in ventral midbrain precursors increases differentiation into dopamine neurons”. eng. In: *Journal of Cell Science* 117.Pt 24 (Nov. 2004), pp. 5731–5737. ISSN: 0021-9533. DOI: 10.1242/jcs.01505.

[329] Ola Awad et al. “Altered Differentiation Potential of Gaucher’s Disease iPSC Neuronal Progenitors due to Wnt/β-Catenin Downregulation”. eng. In: *Stem Cell Reports* 9.6 (Dec. 2017), pp. 1853–1867. ISSN: 2213-6711. DOI: 10.1016/j.stemcr.2017.10.029.

[330] Noel Moya et al. “Endogenous WNT signaling regulates hPSC-derived neural progenitor cell heterogeneity and specifies their regional identity”. eng. In: *Stem Cell Reports* 3.6 (Dec. 2014), pp. 1015–1028. ISSN: 2213-6711. DOI: 10.1016/j.stemcr.2014.10.004.

[331] Olga Momčilović et al. “Genome wide profiling of dopaminergic neurons derived from human embryonic and induced pluripotent stem cells”. eng. In: *Stem Cells and Development* 23.4 (Feb. 2014), pp. 406–420. ISSN: 1557-8534. DOI: 10.1089/scd.2013.0412.

[332] Enrique M. Toledo, Daniel Gyllborg, and Ernest Arenas. “Translation of WNT developmental programs into stem cell replacement strategies for the treatment of Parkinson’s disease”. eng. In: *British Journal of Pharmacology* 174.24 (Dec. 2017), pp. 4716–4724. ISSN: 1476-5381. DOI: 10.1111/bph.13871.

[333] Daniela Berg et al. “MDS research criteria for prodromal Parkinson’s disease”. eng. In: *Movement Disorders: Official Journal of the Movement Disorder Society* 30.12 (Oct. 2015), pp. 1600–1611. ISSN: 1531-8257. DOI: 10.1002/mds.26431.

[334] Kie Itoh et al. “Mitochondrial dynamics in neurodegeneration”. eng. In: *Trends in Cell Biology* 23.2 (Feb. 2013), pp. 64–71. ISSN: 1879-3088. DOI: 10.1016/j.tcb.2012.10.006.

[335] Gyeong Joon Moon, Minsang Shin, and Sang Ryong Kim. “Upregulation of Neuronal Rheb(S16H) for Hippocampal Protection in the Adult Brain”. eng. In: *International Journal of Molecular Sciences* 21.6 (Mar. 2020), p. 2023. ISSN: 1422-0067. DOI: 10.3390/ijms21062023.

[336] Dennis W. Dickson. “Parkinson’s Disease and Parkinsonism: Neuropathology”. In: *Cold Spring Harbor Perspectives in Medicine* 2.8 (Aug. 2012), a009258. ISSN: 2157-1422. DOI: 10.1101/cshperspect.a009258. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3405828/ (visited on 11/23/2022).

[337] Jeong Won Lee et al. “Patients with scans without evidence of dopaminergic deficit (SWEDD) do not have early Parkinson’s disease: Analysis of the PPMI data”. en. In: *PLOS ONE* 16.2 (Feb. 2021). Publisher: Public Library of Science, e0246881. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0246881. URL: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0246881 (visited on 11/15/2022).

[338] Petra Terešak et al. “Regulation of PRKN-independent mitophagy”. In: *Autophagy* 18.1 (), pp. 24–39. ISSN: 1554-8627. DOI: 10.1080/15548627.2021.1888244. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8865282/ (visited on 11/23/2022).

[339] Iara Pérez-Taboada et al. “Diabetes Causes Dysfunctional Dopamine Neurotransmission Favoring Nigrostriatal Degeneration in Mice”. eng. In: *Movement Disorders: Official Journal of the Movement Disorder Society* 35.9 (Sept. 2020), pp. 1636–1648. ISSN: 1531-8257. DOI: 10.1002/mds.28124.

[340] Rudy Luna et al. “A Comprehensive Review of Neuronal Changes in Diabetics”. In: *Cureus* 13.10 (), e19142. ISSN: 2168-8184. DOI: 10.7759/cureus.19142. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8628358/ (visited on 11/23/2022).

[341] A. Schurr and B. M. Rigor. “Brain anaerobic lactate production: a suicide note or a survival kit?” eng. In: *Developmental Neuroscience* 20.4-5 (1998), pp. 348–357. ISSN: 0378-5866. DOI: 10.1159/000017330.

[342] Edoardo Monfrini et al. “VPS13C-associated Parkinson’s disease: Two novel cases and review of the literature”. eng. In: *Parkinsonism & Related Disorders* 94 (Jan. 2022), pp. 37–39. ISSN: 1873-5126. DOI: 10.1016/j.parkreldis.2021.11.031.

[343] Chien-Tai Hong et al. “Insulin Resistance Promotes Parkinson’s Disease through Aberrant Expression of α-Synuclein, Mitochondrial Dysfunction, and Deregulation of the Polo-Like Kinase 2 Signaling”. In: *Cells* 9.3 (Mar. 2020), p. 740. ISSN: 2073-4409. DOI: 10.3390/cells9030740. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7140619/ (visited on 11/23/2022).

[344] Yiting Jia et al. “Targeting macrophage TFEB-14-3-3 epsilon Interface by naringenin inhibits abdominal aortic aneurysm”. en. In: *Cell Discovery* 8.1 (Mar. 2022). Number: 1 Publisher: Nature Publishing Group, pp. 1–21. ISSN: 2056-5968. DOI: 10.1038/s41421-021-00363-1. URL: https://www.nature.com/articles/s41421-021-00363-1 (visited on 11/23/2022).

[345] Rachel Underwood et al. “14-3-3 mitigates alpha-synuclein aggregation and toxicity in the in vivo preformed fibril model”. eng. In: *Acta Neuropathologica Communications* 9.1 (Jan. 2021), p. 13. ISSN: 2051-5960. DOI: 10.1186/s40478-020-01110-5.

[346] Milena Pinto et al. “Pioglitazone ameliorates the phenotype of a novel Parkinson’s disease mouse model by reducing neuroinflammation”. eng. In: *Molecular Neurodegeneration* 11 (Apr. 2016), p. 25. ISSN: 1750-1326. DOI: 10.1186/s13024-016-0090-7.

[347] Lijun Cao et al. “A novel dual GLP-1 and GIP incretin receptor agonist is neuroprotective in a mouse model of Parkinson’s disease by reducing chronic inflammation in the brain”. eng. In: *Neuroreport* 27.6 (Apr. 2016), pp. 384–391. ISSN: 1473-558X. DOI: 10.1097/WNR.0000000000000548.

[348] A. A. Martinez et al. “Activation of PPAR gamma receptors reduces levodopa-induced dyskinesias in 6-OHDA-lesioned rats”. eng. In: *Neurobiology of Disease* 74 (Feb. 2015), pp. 295–304. ISSN: 1095-953X. DOI: 10.1016/j.nbd.2014.11.024.

[349] Eulàlia Segur-Bailach et al. “Over-Mutated Mitochondrial, Lysosomal and TFEB-Regulated Genes in Parkinson’s Disease”. In: *Journal of Clinical Medicine* 11.6 (Mar. 2022), p. 1749. ISSN: 2077-0383. DOI: 10.3390/jcm11061749. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8951534/ (visited on 11/23/2022).

[350] Lyudmila F. Gulyaeva and Nicolay E. Kushlinskiy. “Regulatory mechanisms of microRNA expression”. In: *Journal of Translational Medicine* 14.1 (May 2016), p. 143. ISSN: 1479-5876. DOI: 10.1186/s12967-016-0893-x. URL: https://doi.org/10.1186/s12967-016-0893-x (visited on 01/01/2023).

[351] A. Garg. “Modeling stochasticity and robustness in gene regulatory networks”. en. In: *Bioinforma. Oxf. Engl* 25 (2009), pp. 101–109.

[352] D. Murrugarra. “Modeling stochasticity and variability in gene regulatory networks”. en. In: *EURASIP J. Bioinforma. Syst. Biol* (2012), p. 5.

[353] M. Koltai. “Exact solving and sensitivity analysis of stochastic continuous time Boolean models”. en. In: *BMC Bioinformatics* 21 (2020), p. 241.

[354] Fabian Fröhlich et al. “Parameter estimation for dynamical systems with discrete events and logical operations”. eng. In: *Bioinformatics (Oxford, England)* 33.7 (Apr. 2017), pp. 1049–1056. ISSN: 1367-4811. DOI: 10.1093/bioinformatics/btw764.

[355] Fabian Fröhlich et al. “Scalable Parameter Estimation for Genome-Scale Biochemical Reaction Networks”. eng. In: *PLoS computational biology* 13.1 (Jan. 2017), e1005331. ISSN: 1553-7358. DOI: 10.1371/journal.pcbi.1005331.

[356] Miroslav Kratochvíl et al. “COBREXA.jl: constraint-based reconstruction and exascale analysis”. In: *Bioinformatics* 38.4 (Feb. 2022), pp. 1171–1172. ISSN: 1367-4803. DOI: 10.1093/bioinformatics/btab782. URL: https://doi.org/10.1093/bioinformatics/btab782 (visited on 01/08/2023).

[357] Fabian Fröhlich et al. “Efficient Parameter Estimation Enables the Prediction of Drug Response Using a Mechanistic Pan-Cancer Pathway Model”. eng. In: *Cell Systems* 7.6 (Dec. 2018), 567–579.e6. ISSN: 2405-4712. DOI: 10.1016/j.cels.2018.10.013.

[358] J. L. Snoep and B. G. Olivier. “Java Web Simulation (JWS); a web based database of kinetic models”. eng. In: *Molecular Biology Reports* 29.1-2 (2002), pp. 259–263. ISSN: 0301-4851. DOI: 10.1023/a:1020350518131.

[359] K. Tiwari. “Reproducibility in systems biology modelling”. cs. In: *Mol. Syst. Biol* 17 (2021).

[360] A. Niarakis. “Setting the basis of best practices and standards for curation and annotation of logical models in biology—highlights of the [BC]2 2019 CoLoMoTo/SysMod Workshop”. en. In: *Brief. Bioinform* 22 (2021), pp. 1848–1859.

Appendix A

Supplementary data and information

A.1 Source code and data repository

All code and data related to this thesis is stored in a gitlab repository at [183]. The overall structure of the Gitlab repo is illustrated in Figure A.1.

A.2 The overall research process

The high-level approach to Boolean model construction is summarized in Algorithm 1.

The general methodology scheme used for analysing the models is summarized in Algorithm 2.

The algorithm for integrating the Boolean models with omics data is summarized in Algorithm 3.

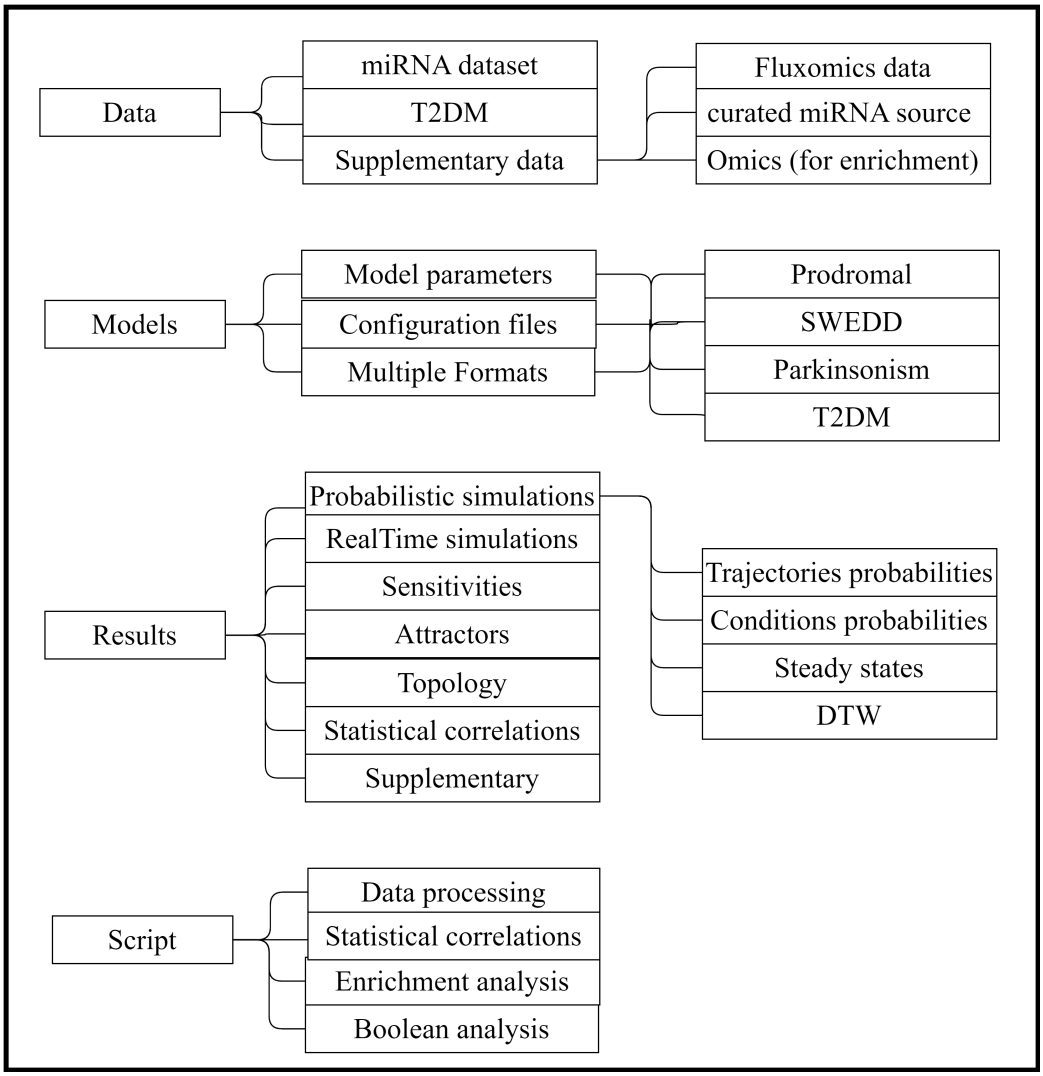


Figure A.1 The structure of the Gitlab repository

Algorithm 1 Construction of Boolean models

```
1: procedure PD map selection
2:     Identify relevant PD map diagrams for biological system of interest
3:     Filter diagrams for relevancy and remove duplicates
4:     Sort remaining diagrams by level of detail and complexity
5: end procedure
6: procedure Translation with CaSQ
7:     Load CaSQ software
8:     for each diagram in sorted set of diagrams do
9:         Select diagram
10:        Follow CaSQ prompts to translate diagram into Boolean model
11:    end for
12:    Models in SBML-qual
13: end procedure
14: procedure Model correctness
15:    for each Boolean model do
16:        Validate model correctness by comparing predictions to experimental
    data
17:        if model is incorrect then
18:            Revise the Boolean functions and their relevance
19:        end if
20:    end for
21: end procedure
```

Algorithm 2 Model analysis

```
1: procedure Model updating schemes
2:     Choose synchronous or asynchronous updating scheme
3: end procedure
4: procedure Attractor search
5:     Identify stable states (attractors) in the model
6:     for each transition between attractors do
7:         Identify transition
8:         Compare transition to experimental data
9:     end for
10: end procedure
11: procedure Perturbation analysis
12:     Identify set of inputs (perturbations) to the model
13:     for each input do
14:         Simulate model with input
15:         Compare output changes to experimental data
16:     end for
17: end procedure
```

Algorithm 3 Integration of Boolean models with omics data

```
1: procedure PPMI-miRNA dataset
2:     Load PPMI dataset
3:     Identify set of miRNAs in dataset
4:     for each miRNA in set do
5:         Identify targets of miRNA using prediction tool
6:     end for
7: end procedure
8: procedure Type two Diabetes mellitus dataset
9:     Load Type two Diabetes mellitus dataset
10:    Identify set of relevant variables in dataset
11:    Update Boolean model using variables
12: end procedure
13: procedure Stochastic Boolean model simulation
14:    Use stochastic simulation tool to simulate Boolean model
15:    Compare model behavior to experimental data
16: end procedure
```

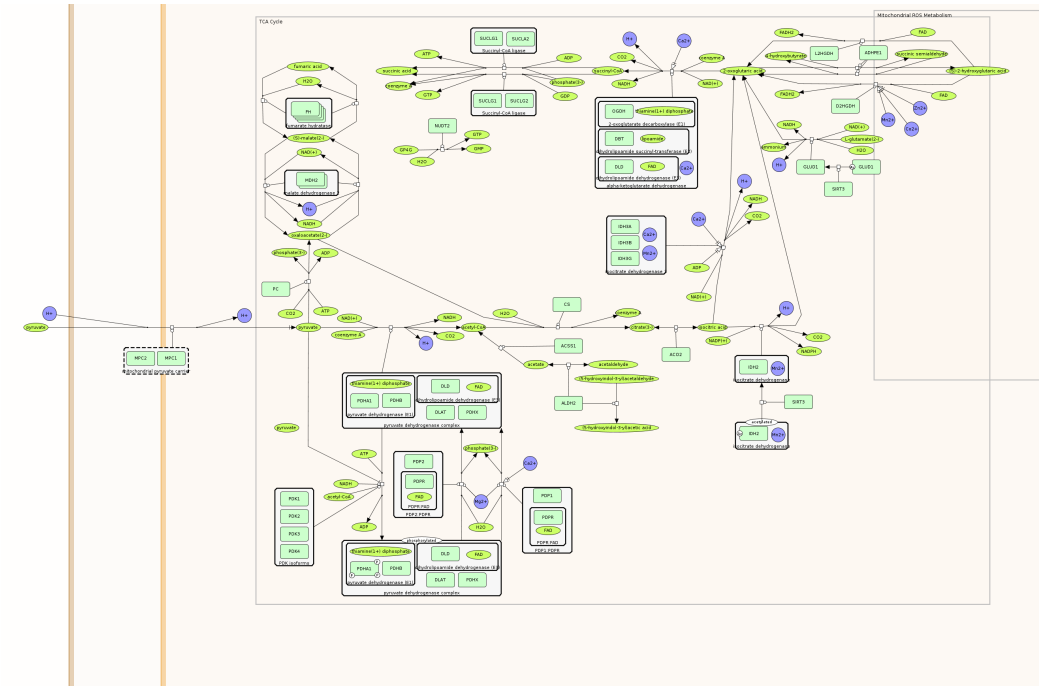


Figure A.2 Dopamine transcription pathway

A.3 Selected diagrams from PD map

For completeness, we include the following pathway diagrams from PD map:

- Dopamine transcription pathway (Figure A.2)
- FOXO3 activity pathway (Figure A.3)
- PI3KAKT signalling pathway (Figure A.4)
- mTOR pathway (Figure A.5)
- PRKN signalling (Figure A.6)
- TCA cycle (Figure A.7)

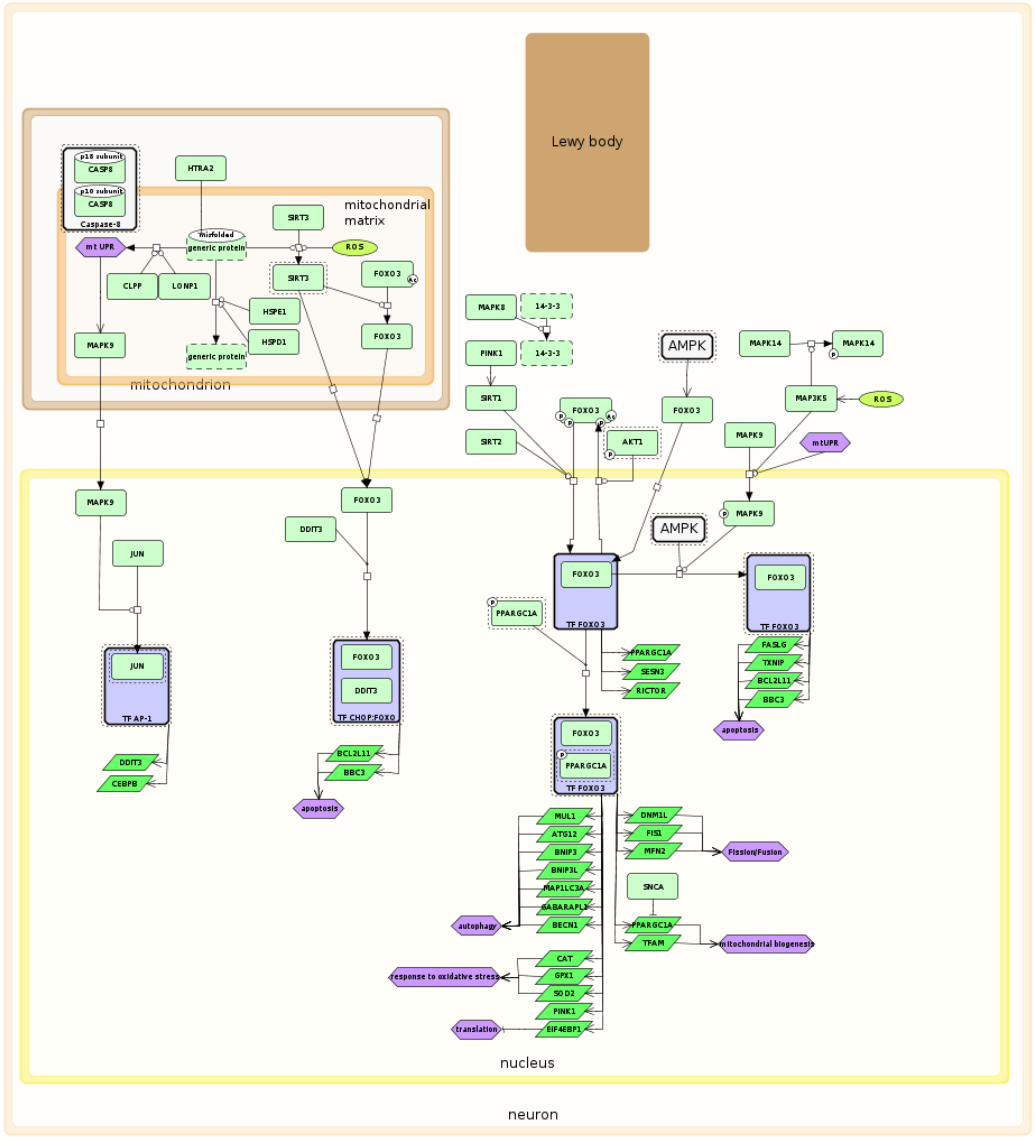


Figure A.3 FOXO3 Activity pathway

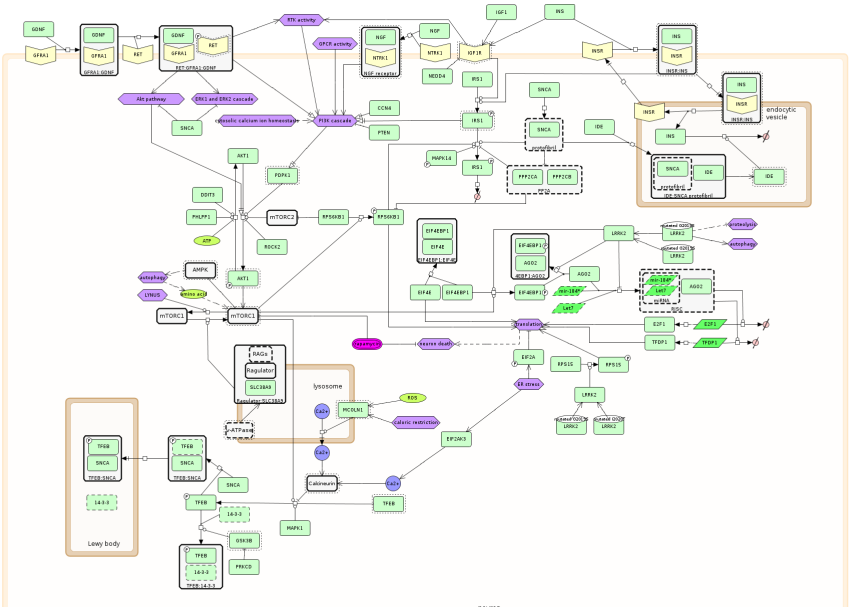


Figure A.4 PI3KAKT signalling pathway

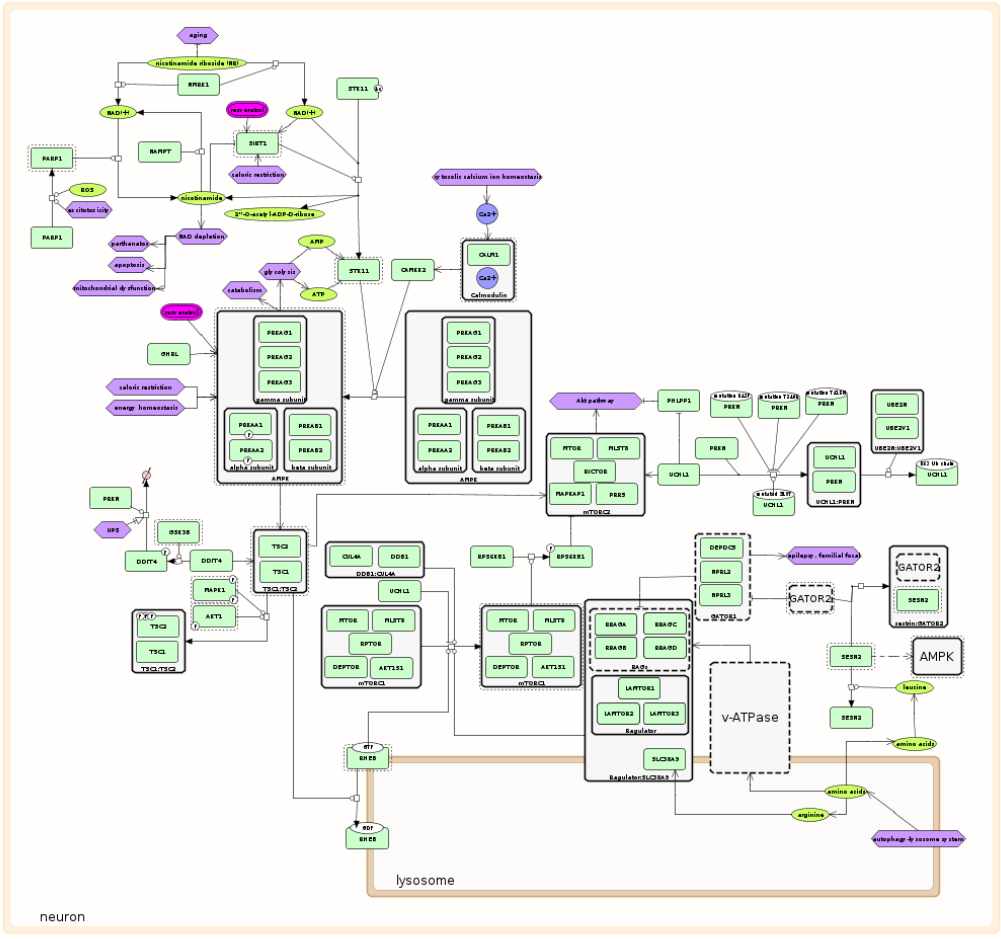


Figure A.5 mTOR pathway

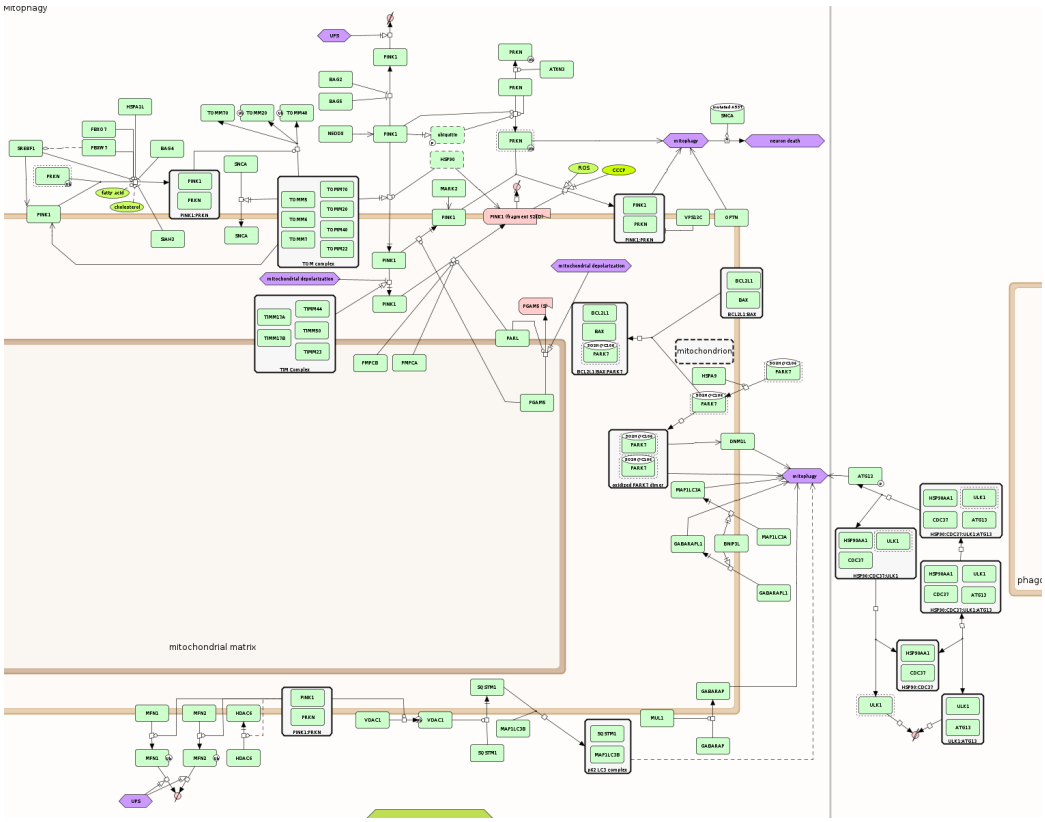


Figure A.6 PRKN signalling pathway

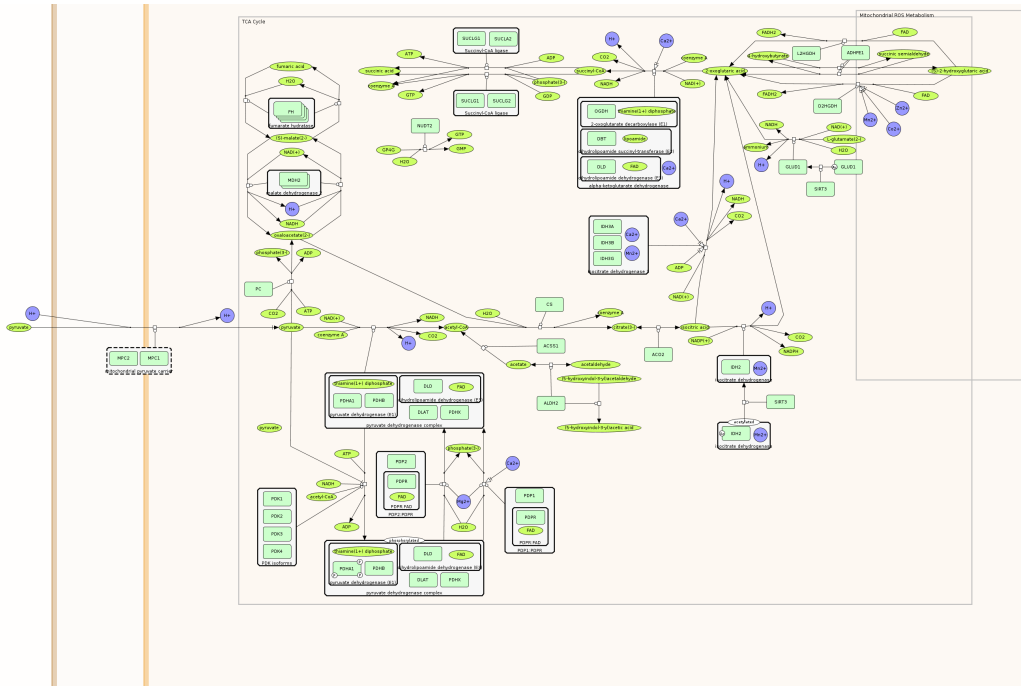


Figure A.7 TCA cycle

Appendix B

Selected simulation result details

B.1 The simulation graphs from Boolean simulations

This section details the simulation graphs that resulted from Boolean models and Probabilistic BMs simulation. All simulation graphs are also available in the thesis data repository (see Appendix A.1).

The examples from simulation graphs from CellCollective platform are shown in Figures B.1 to B.4. Interactive demo and guidance on how to construct and simulate the models is available at `https://cellcollective.org/`.

Examples from simulation graphs from pyMaBoSS are shown in Figures B.5 to B.18.

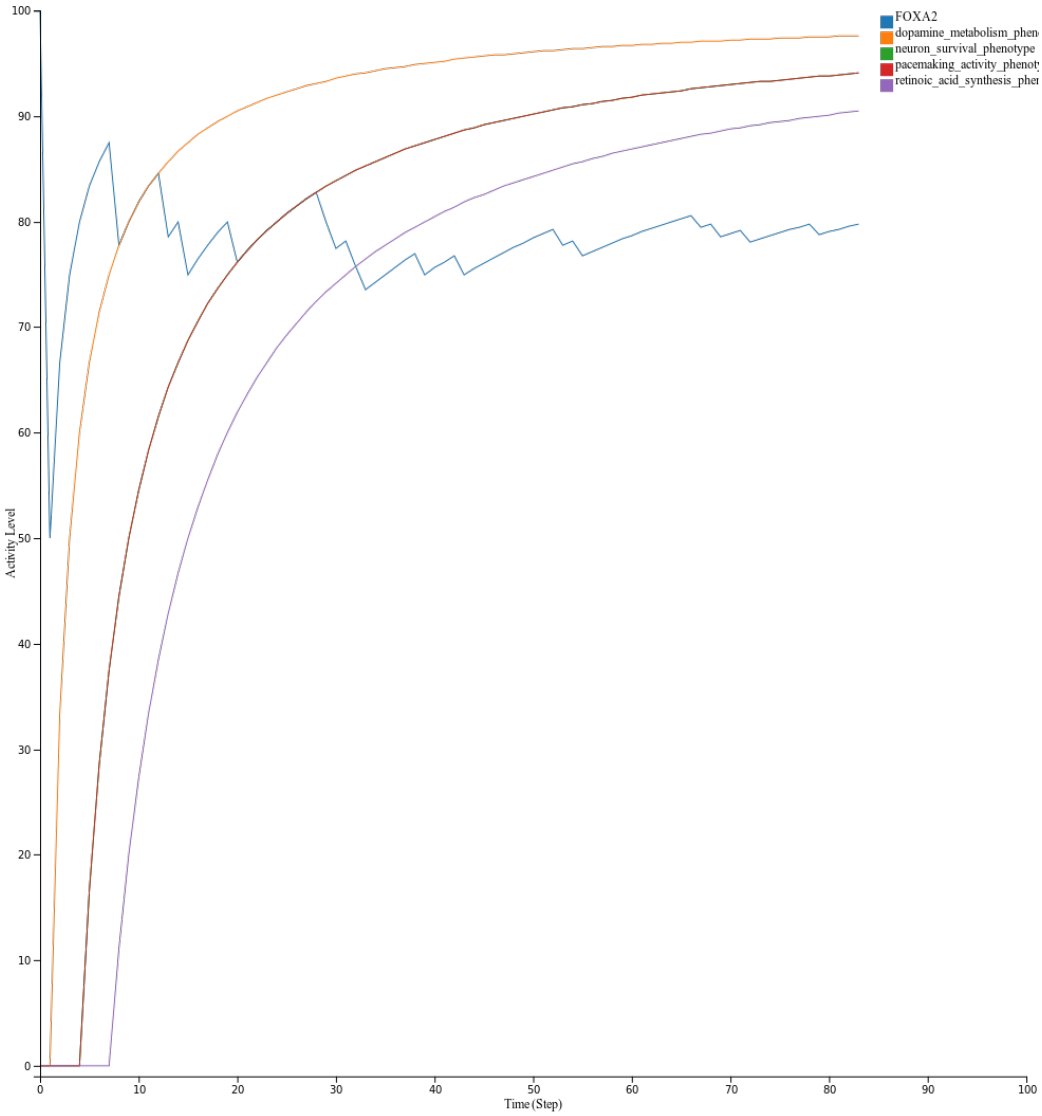


Figure B.1 Dopamine transcription simulation

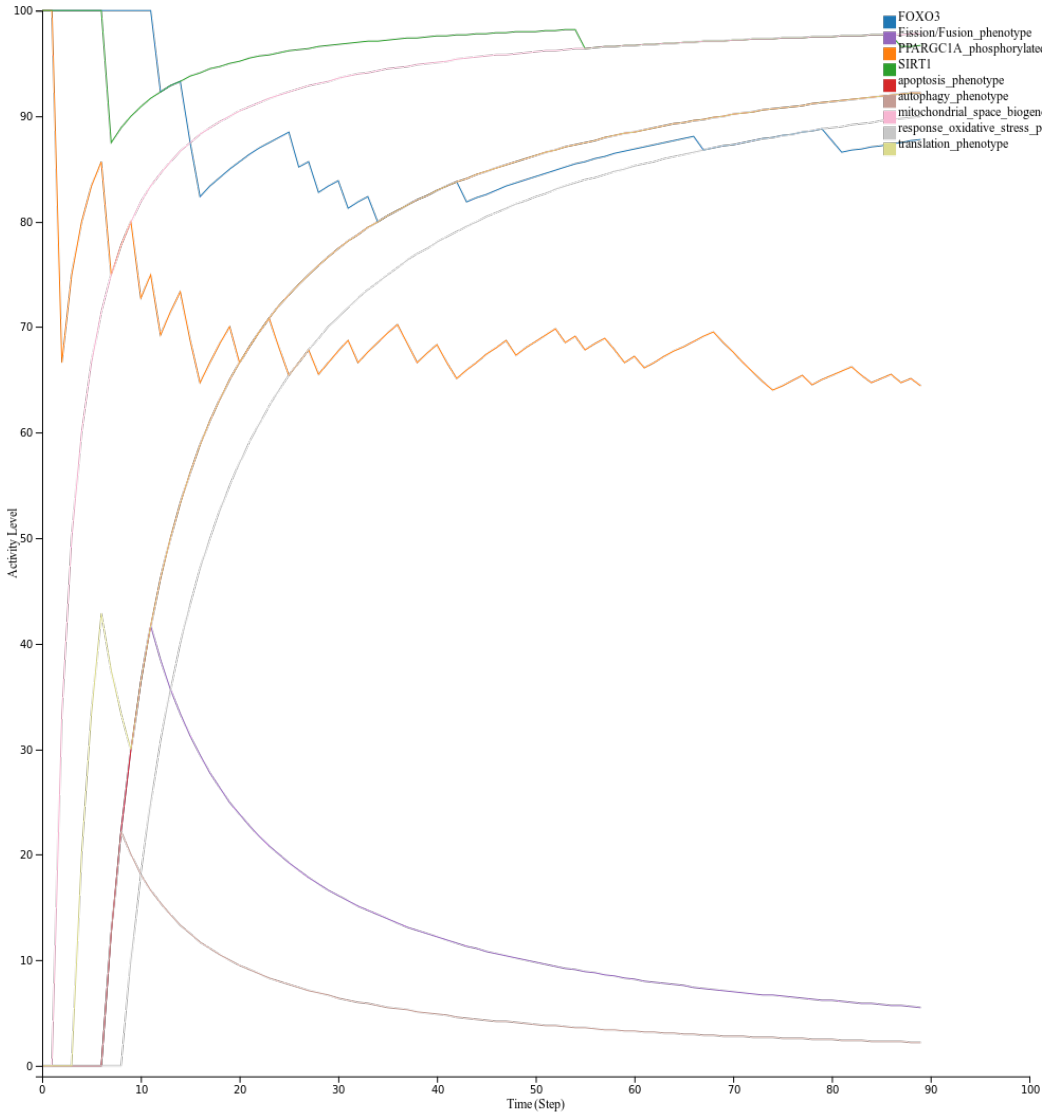


Figure B.2 FOXO3 simulation

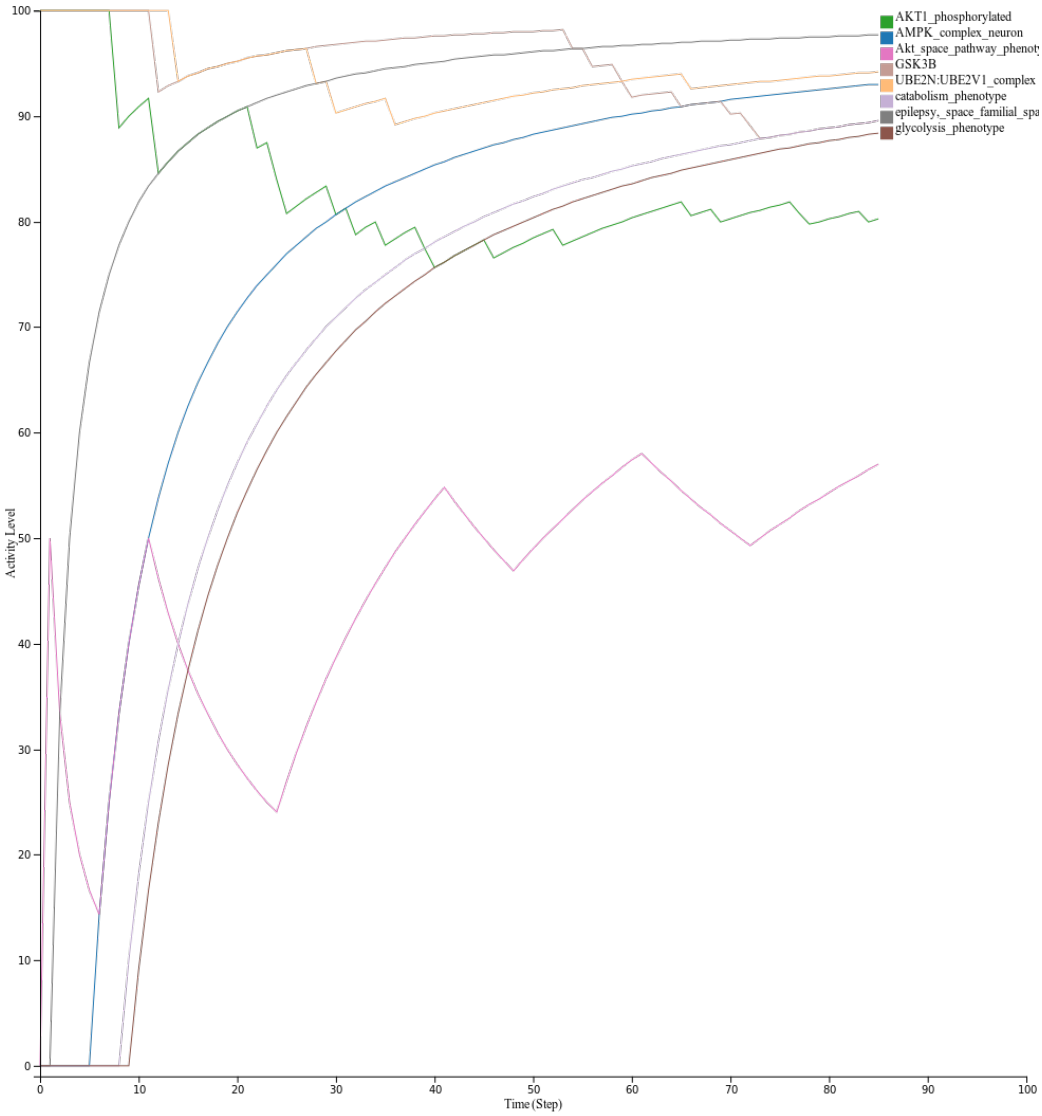


Figure B.3 mTOR simulation

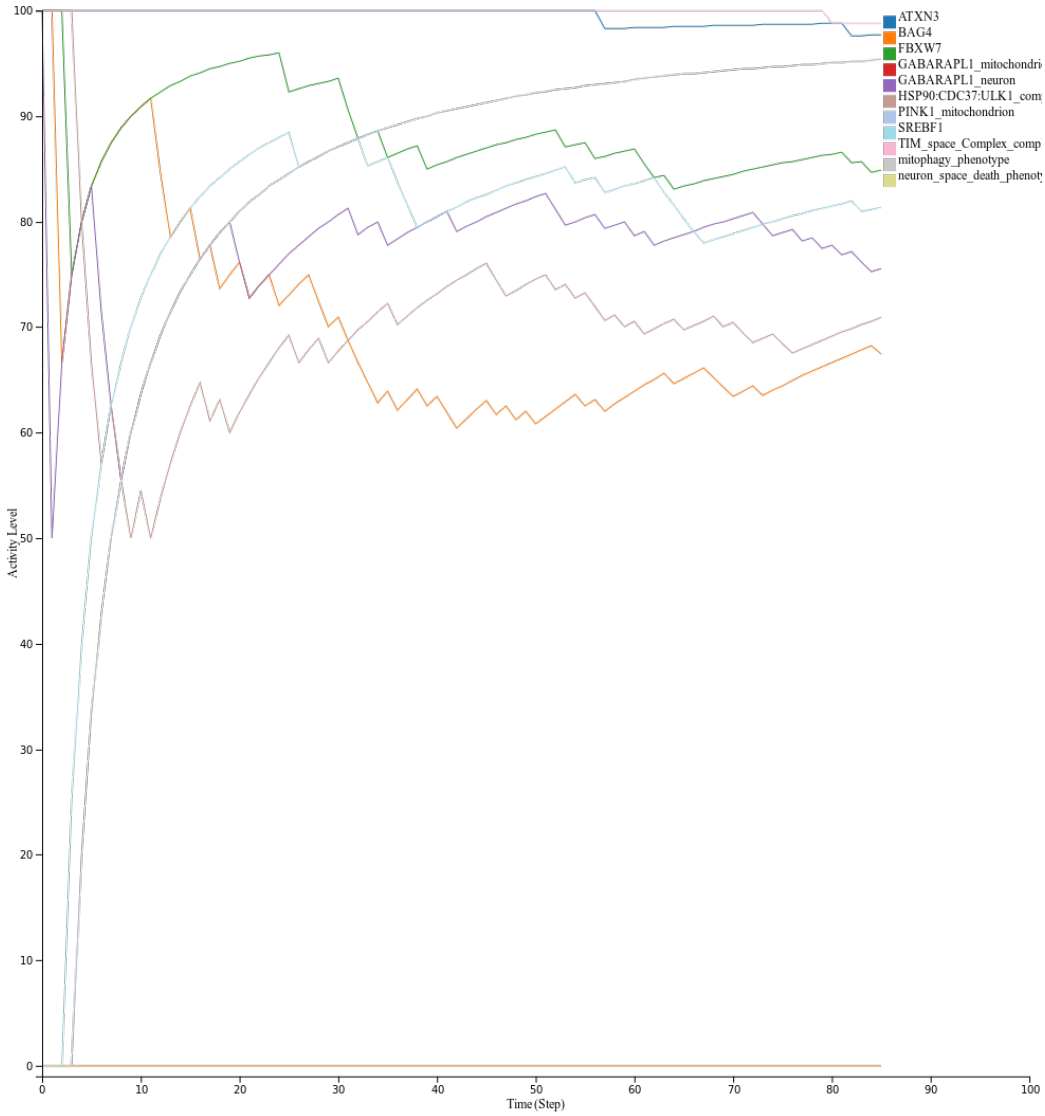


Figure B.4 PRKN simulation

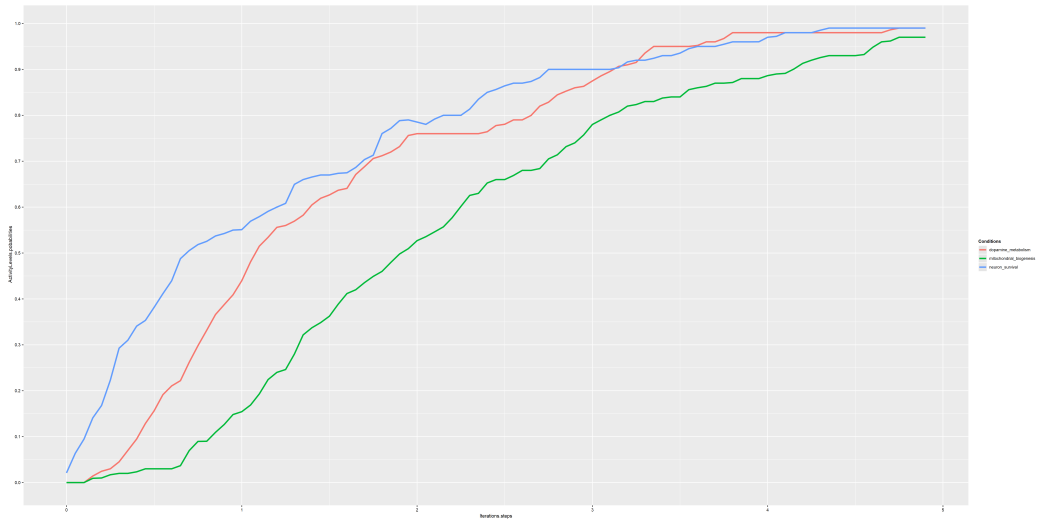


Figure B.5 Dopamine transcription simulation in the prodromal

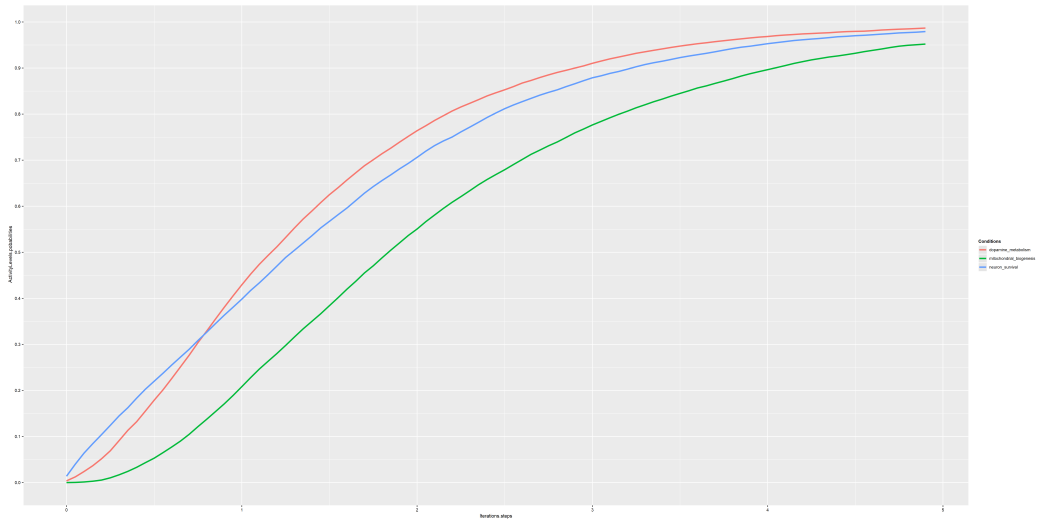


Figure B.6 Dopamine transcription model in the SWEDD

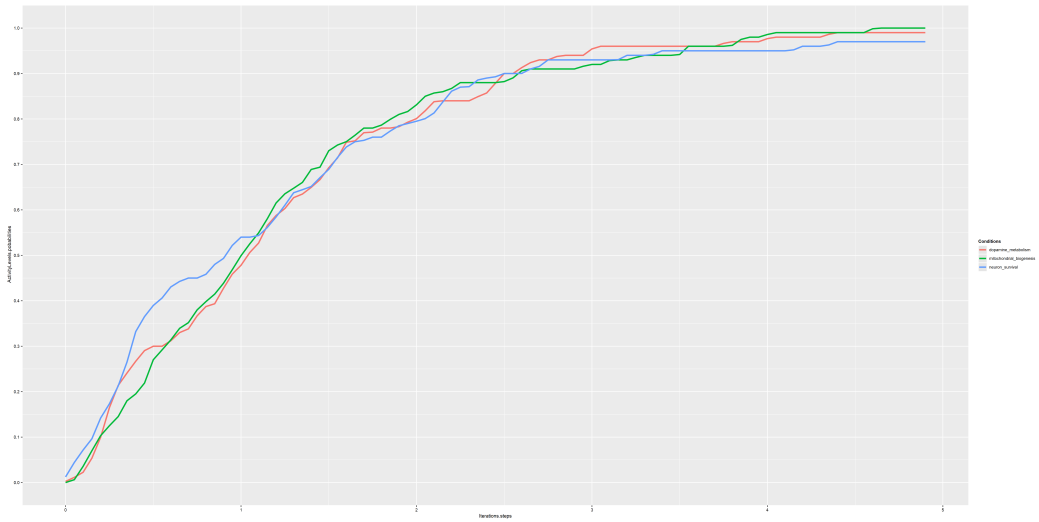


Figure B.7 Dopamine transcription model in the Parkinsonism

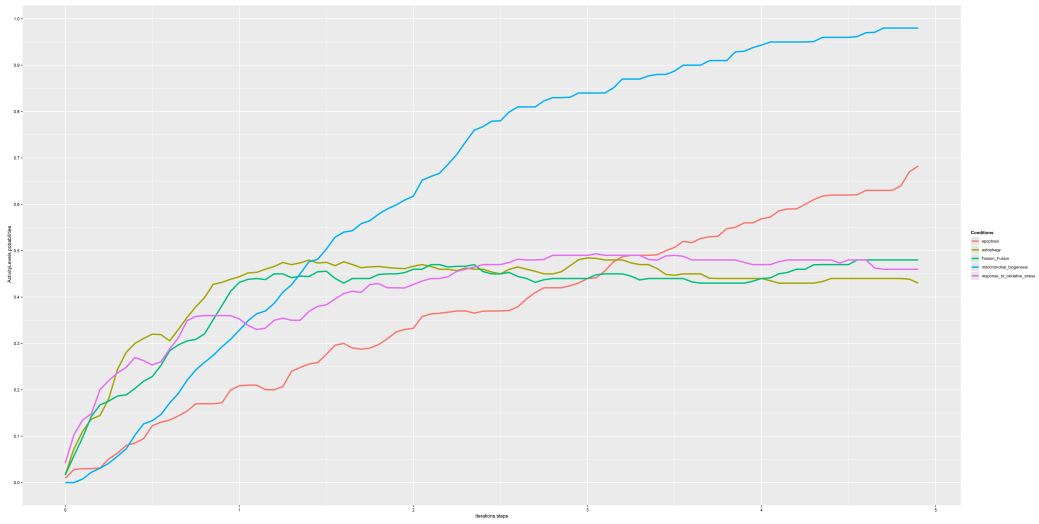


Figure B.8 FOXO3 Activity simulation in prodromal

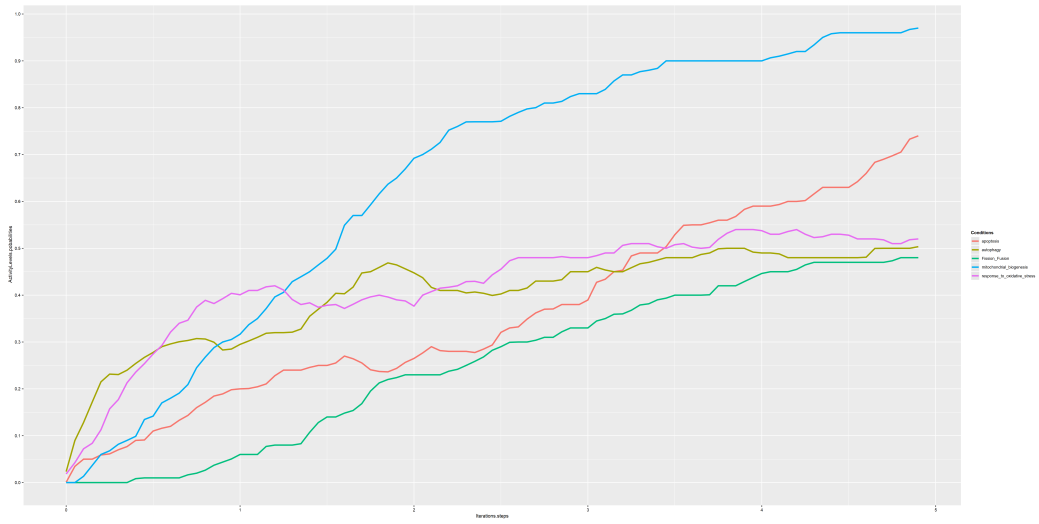


Figure B.9 FOXO3 Activity simulation in SWEDD

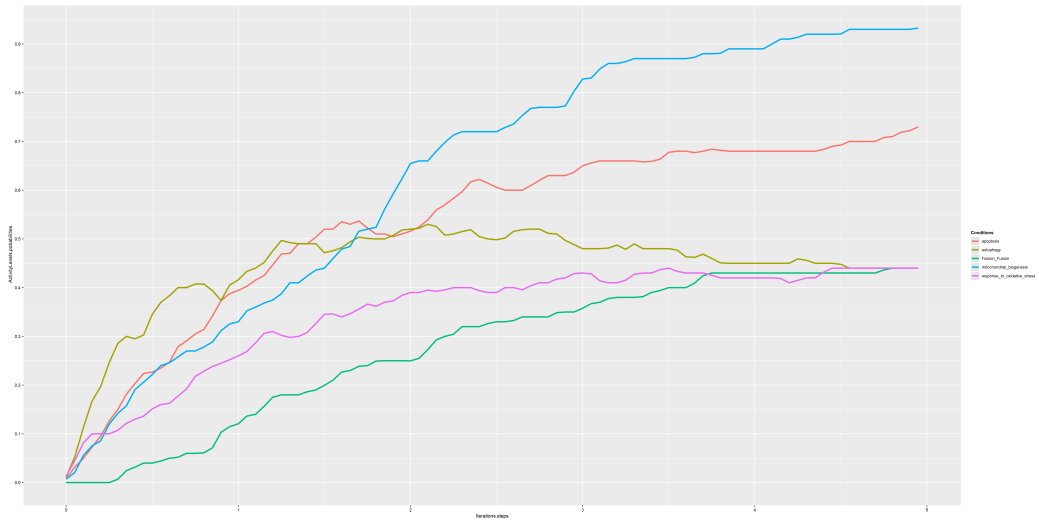


Figure B.10 FOXO3 Activity simulation in parkinsonism

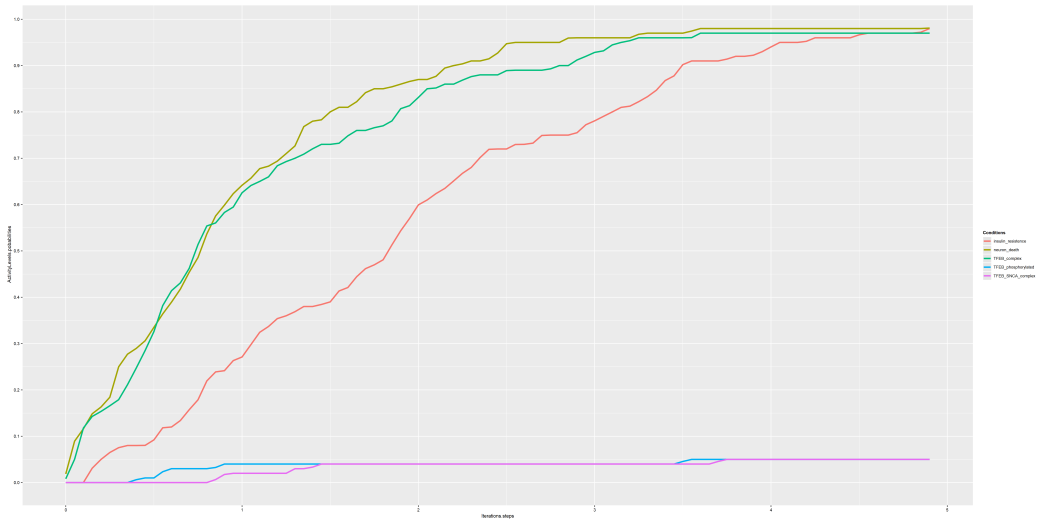


Figure B.11 PI3KAKT model in prodromal

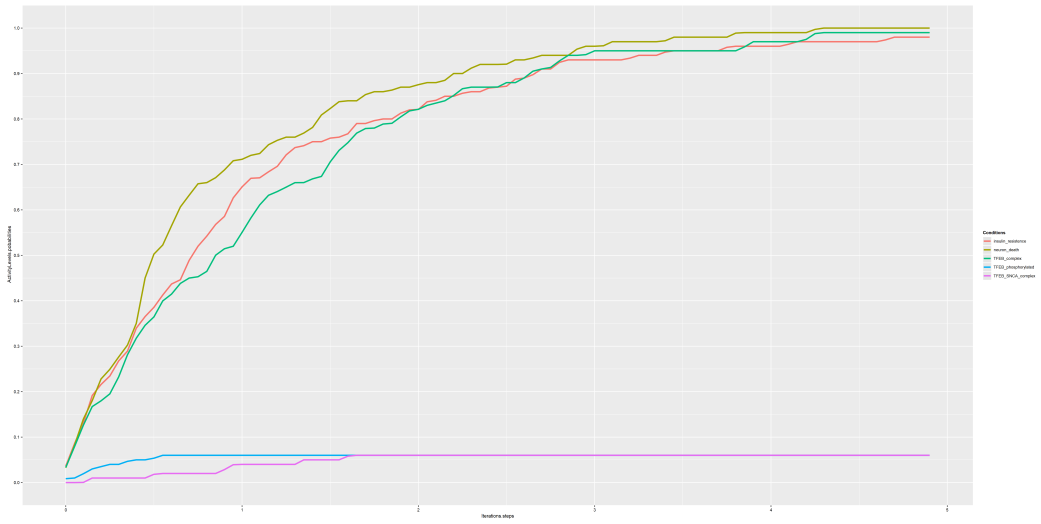


Figure B.12 PI3KAKT model in SWEDD

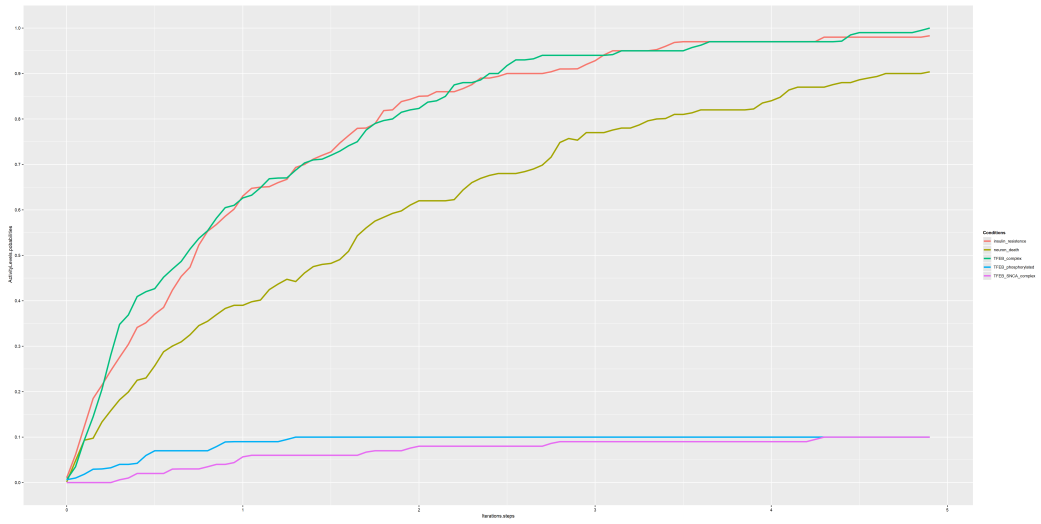


Figure B.13 PI3KAKT model in parkinsonism

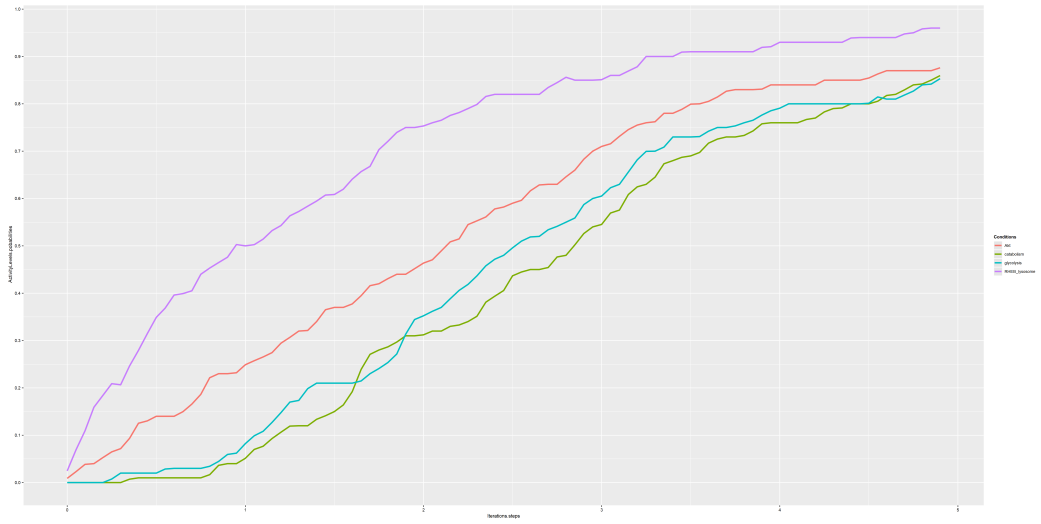


Figure B.14 mTOR model in prodromal

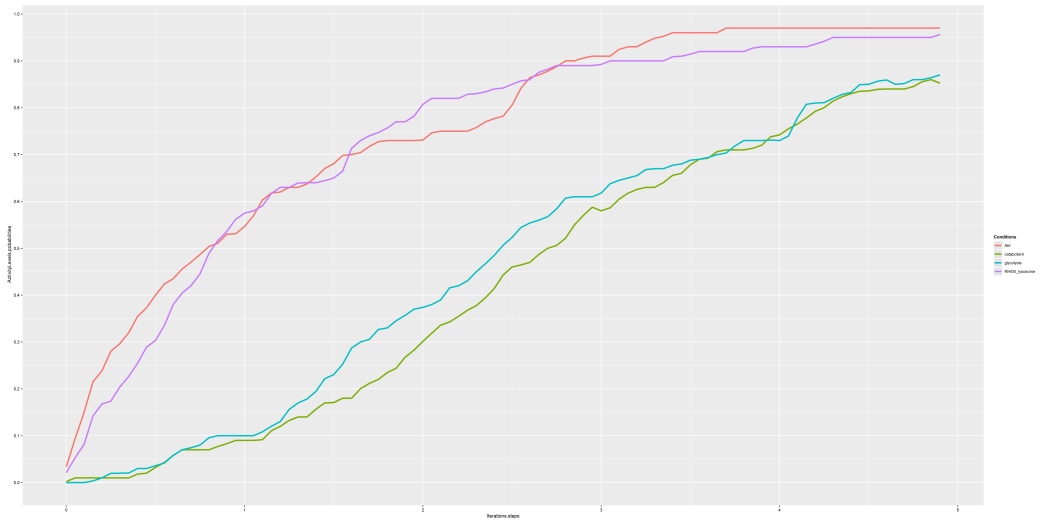


Figure B.15 mTOR model in SWEDD

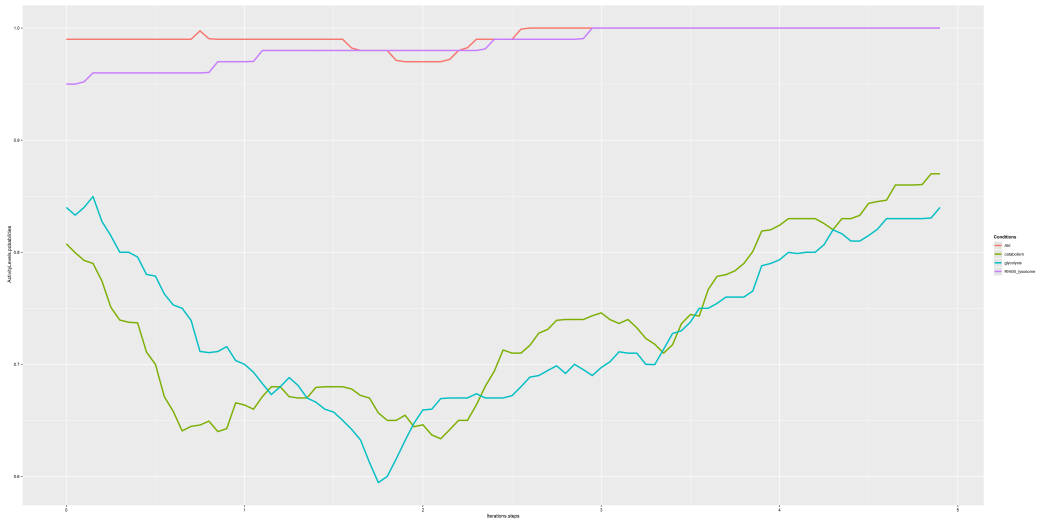


Figure B.16 mTOR model in SWEDD

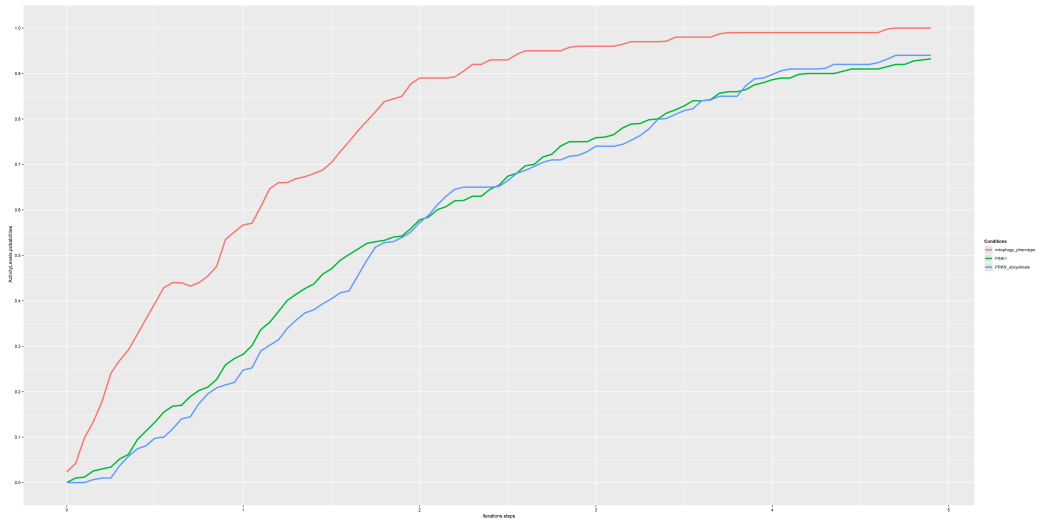


Figure B.17 PRKN model in prodromal and SWEDD

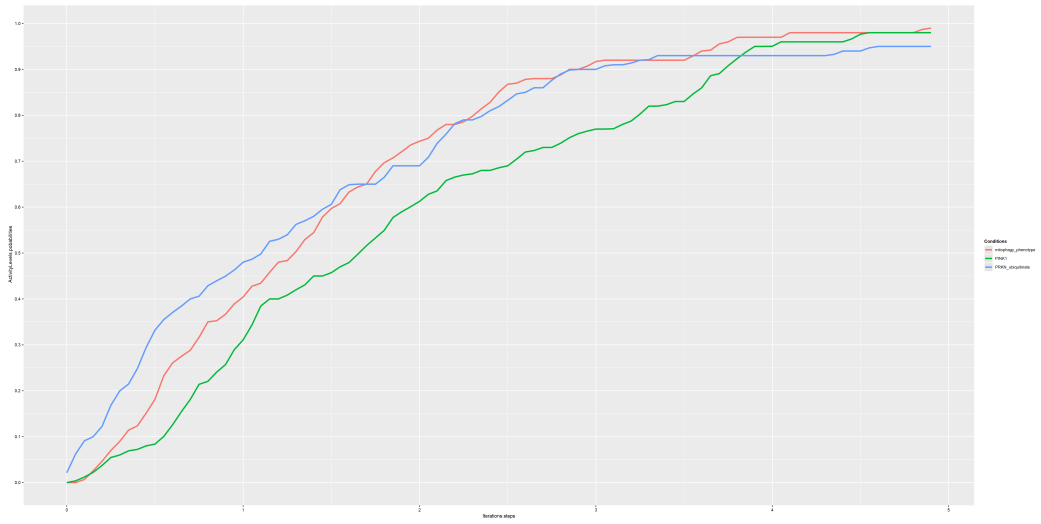


Figure B.18 PRKN model in parkinsonism

B.2 The sensitivity analysis in response to mutations

Distances computed between the original and perturbed attractors are summarized in Table B.1 for knockouts and Table B.2 for overexpressions.

Pathway	Group ID	Identity-based distance	Similarity-based distance
Pi3k/akt	RPS6KB1	0.582524272	0.012230182
	PHLPP1	0.582524272	0.007517202
	WNT1	0.563106796	0.013361297
	WNT3	0.553398058	0.008483363
	PRKN	0.553398058	0.005372797
TCA cycle	PDP2:PDPR_complex	0.630769231	0.010729783
	alpha-ketoglutaratedehydrogenase_complex	0.630769231	0.009704142
	PDP1:PDPR_complex	0.6	0.01025641
	SIRT3	0.6	0.037790927
	isocitratedehydrogenase_complex	0.6	0.012544379
	GLUD1	0.6	0.01530572
PRKN	SREBF1	0.648148148	0.012003
	FBXW7	0.648148148	0.024005
	OPTN	0.648148148	0.012003
	HSP90	0.611111111	0.012003
PPARGC1A	IDH3G_rna	0.582089552	0.008539393
	TF_YY1_complex	0.582089552	0.060741071
	IDH3A_rna	0.582089552	0.008539393
	NDUFS8_rna	0.582089552	0.008539393
	ATP5MC1_rna	0.582089552	0.008539393
mTOR	TSC1:TSC2_complex_neuron	0.603174603	0.0095742
	SESN2	0.603174603	0.0095742
	NAMPT	0.587301587	0.009322247
	ROS	0.571428571	0.009070295
	Akt	0.571428571	0.009070295
Foxo3	MAP3K5	0.61971831	0.008728427
	EIF4EBP1_rna	0.605633803	0.008530054
	ATG12_rna	0.605633803	0.008530054
	BECN1_rna	0.605633803	0.008530054
	BBC3_rna	0.591549296	0.00833168
	JUN	0.577464789	0.008133307
	SIRT1	0.577464789	0.008133307
Dopamine transcription	EN1	0.852941176	0.013327206
	LRRK2	0.573529412	0.046243107
	FOXA2	0.573529412	0.008961397
	SNCA	0.558823529	0.022575827
	SFPQ	0.529411765	0.008272059
	PIN1	0.514705882	0.008042279
	RXRA	0.5	0.0078125

Table B.1 Examples shows the significant distances between the original and perturbed attractors (Knockouts)

Pathway	Group ID	Identity-based distance	Similarity-based distance
Pi3k/akt	CTNNB1_phosphorylated	1	0.015081535
	EIF4EBP1_phosphorylated	1	0.017862192
	IRS1_phosphorylated	1	0.009708738
	CTNNB1_ubiquitinated_phosphorylated	1	0.009708738
	PDPK1	1	0.014209633
	RPS6KB1_phosphorylated	1	0.010085776
	AKT1_phosphorylated	0.990291262	0.018828353
	TFEB_complex	0.980582524	0.009685173
TCA cycle	2-oxoglutaricacid	1	0.035108481
	hydroxyglutaric_acid	1	0.033609467
	succinic_semialdehyde	1	0.025325444
	acetyl-CoA	1	0.016094675
	oxaloacetate_2	1	0.015384615
	succinyl-CoA	1	0.015384615
	succinic_acid	1	0.015384615
	GMP	0.984615385	0.014674556
	GTP	0.984615385	0.014674556
PRKN	PRKN_ubiquitinated	0.981481481	0.021604938
	PINK1_neuron	0.944444444	0.043552812
	ubiquitin_phosphorylated	0.944444444	0.017489712
	PGAM5__S_	0.907407407	0.016803841
	PINK1_mitochondrion	0.888888889	0.017489712
	PINK1	0.87037037	0.01611797
PPARGC1A	COX5A_rna	0.985074627	0.014702606
	COX7A2_rna	0.985074627	0.014702606
	COX5B_rna	0.985074627	0.014702606
	CYCS_rna	0.985074627	0.015296651
	complex_IV_complex	0.985074627	0.014702606
	SDHB_rna	0.985074627	0.014702606
mTOR	CAMKK2	0.634920635	0.010078105
	MAPK1_phosphorylated_phosphorylated	0.603174603	0.0095742
	AMPK_complex_neuron	0.571428571	0.009070295
	DDB1:CUL4A_complex	0.571428571	0.009070295
	PRKN_neuron	0.555555556	0.008818342
FOXO3	FASLG_rna	0.661971831	0.009323547
	MAPK9_phosphorylated	0.647887324	0.009125174
	FOXO3_neuron	0.591549296	0.00833168
	PPARGC1A_rna	0.591549296	0.00833168
	FOXO3_acetylated_phosphorylated	0.577464789	0.008133307
	BCL2L11_rna	0.577464789	0.008133307
	FIS1_rna	0.577464789	0.008133307
Dopamine transcription	ALDH1A1_rna	1	0.03125
	DRD2_rna	1	0.015625
	TH_rna	1	0.015625
	BDNF_rna	0.985294118	0.015395221
	DDC_rna	0.970588235	0.014820772
	SLC18A2_rna	0.941176471	0.014705882
	SLC6A3_rna	0.941176471	0.014705882
	PITX3	0.911764706	0.026711857
	TF_NR4A2_complex	0.897058824	0.242704504

Table B.2 Examples shows the significant distances between the original and perturbed attractors (overexpressions)

B.3 Attractor cycles representations

The landscape of the network state transitions along with attractor cycles were identified. The returned transition network object has same structures with the normal network object. The transition network is written as a SIF file. The SIF file could be loaded to Cytoscape with the following the steps in Figure B.19.

The result is shown in Figure B.20.

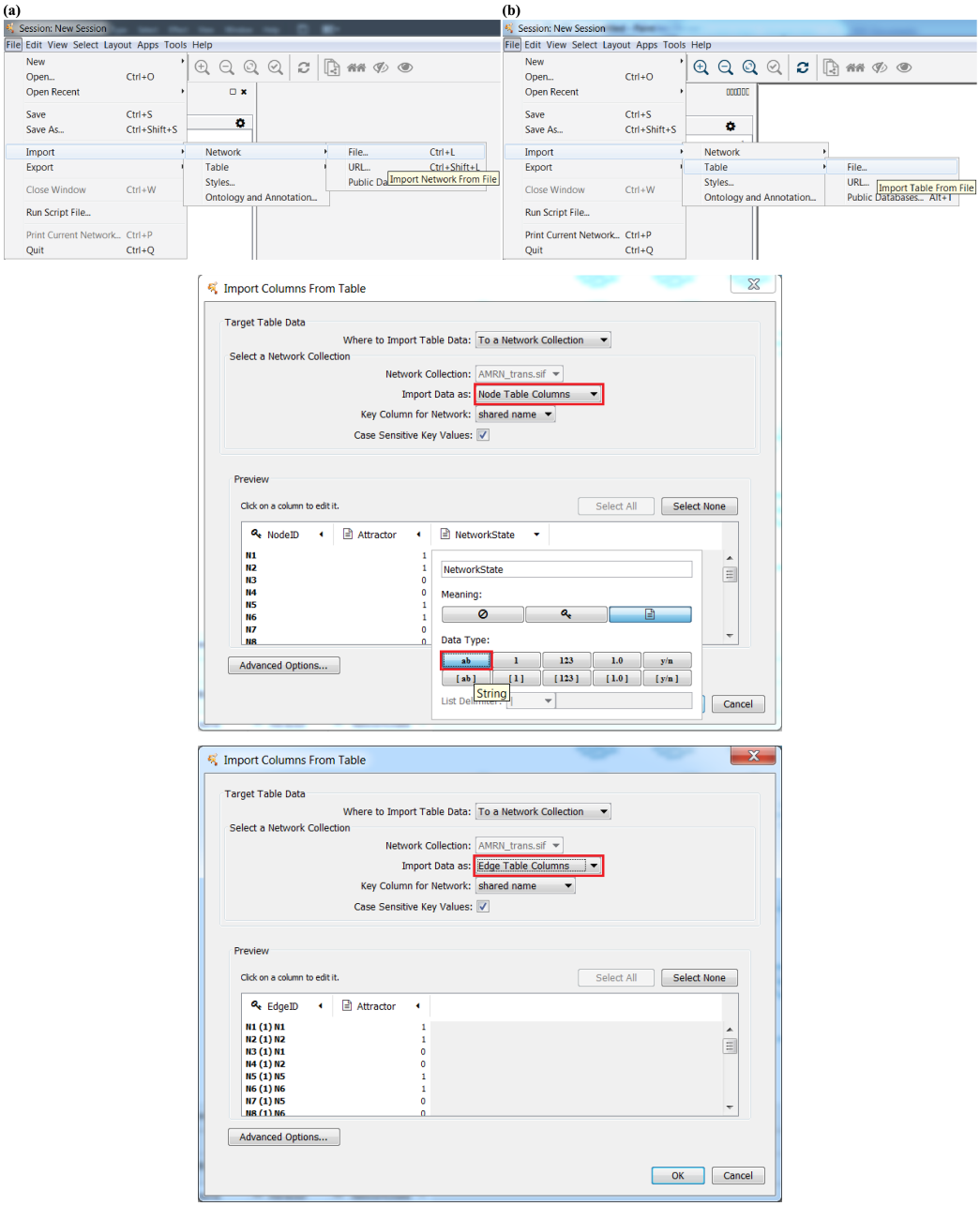


Figure B.19 Steps for loading a SIF fyle in Cytoscape (top to bottom: opening the import dialogs, importing the nodes, importing the edges).

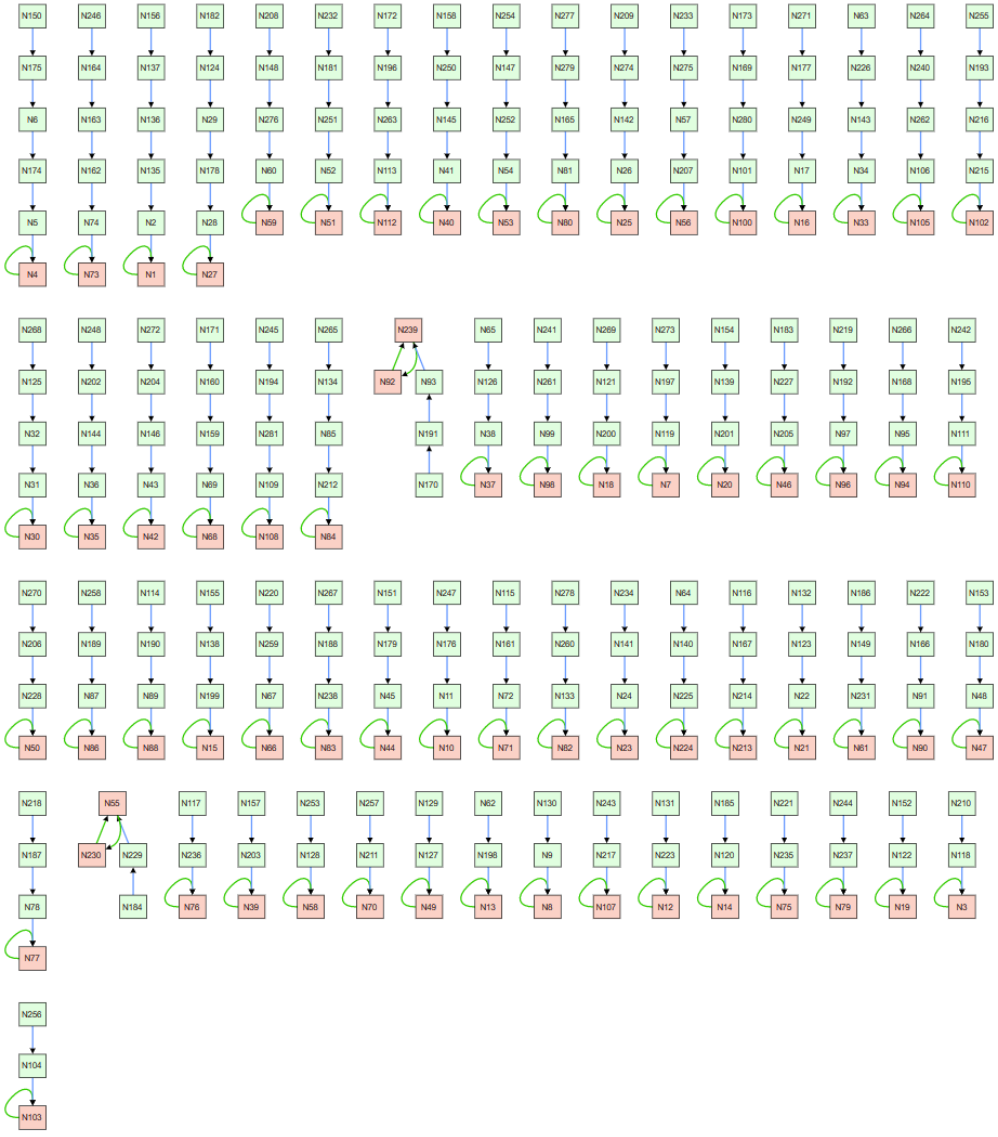


Figure B.20 PPARGC1A pathway attractor

B.4 Stratification related data

An overview of stratification related processes, including enrichment analysis workflow and statistical correlations are shown in Figures B.21 and B.22. Parameters are listed in Table B.4.

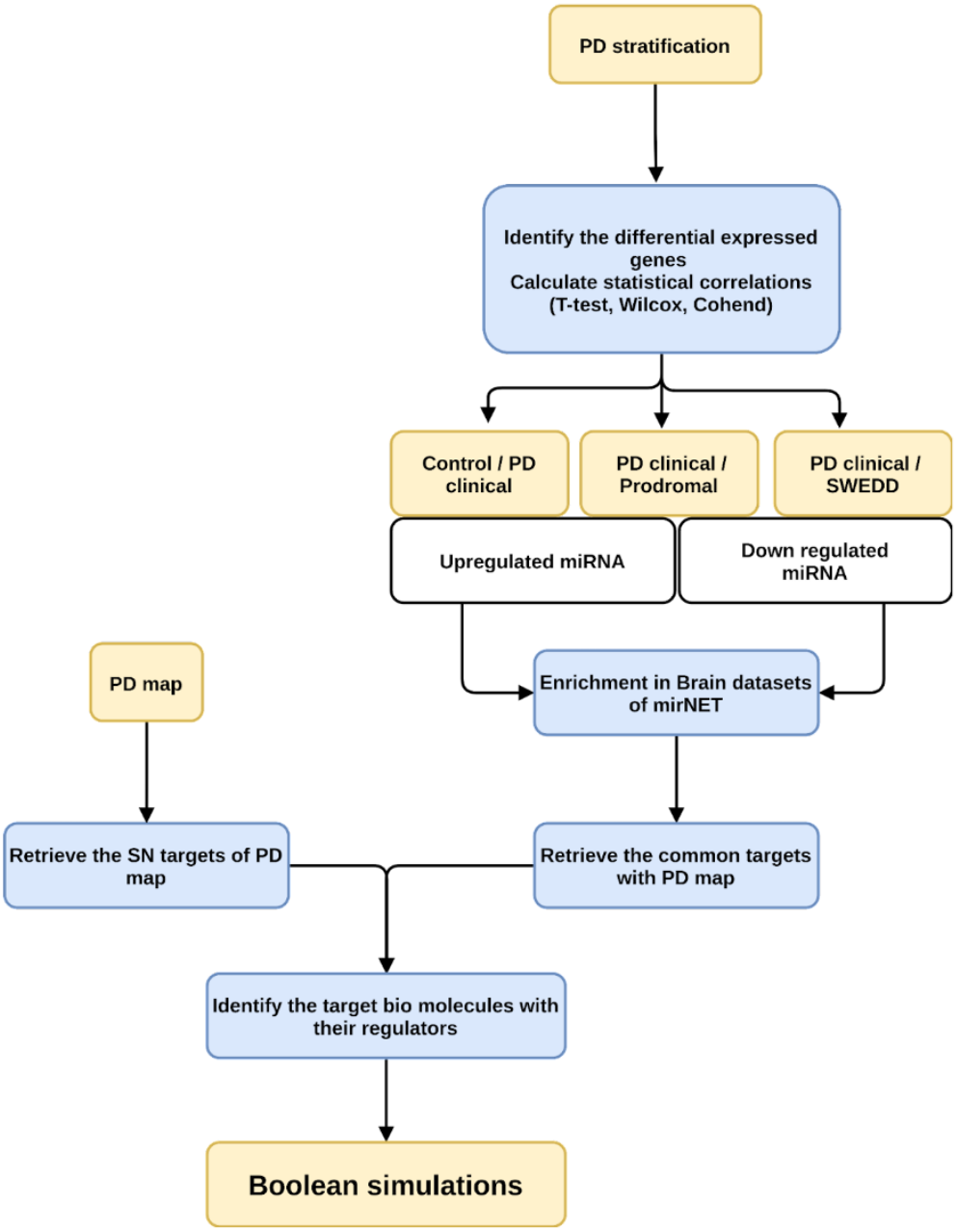


Figure B.21 The stratification process with enrichment analysis pipeline to define the common targets with PD map

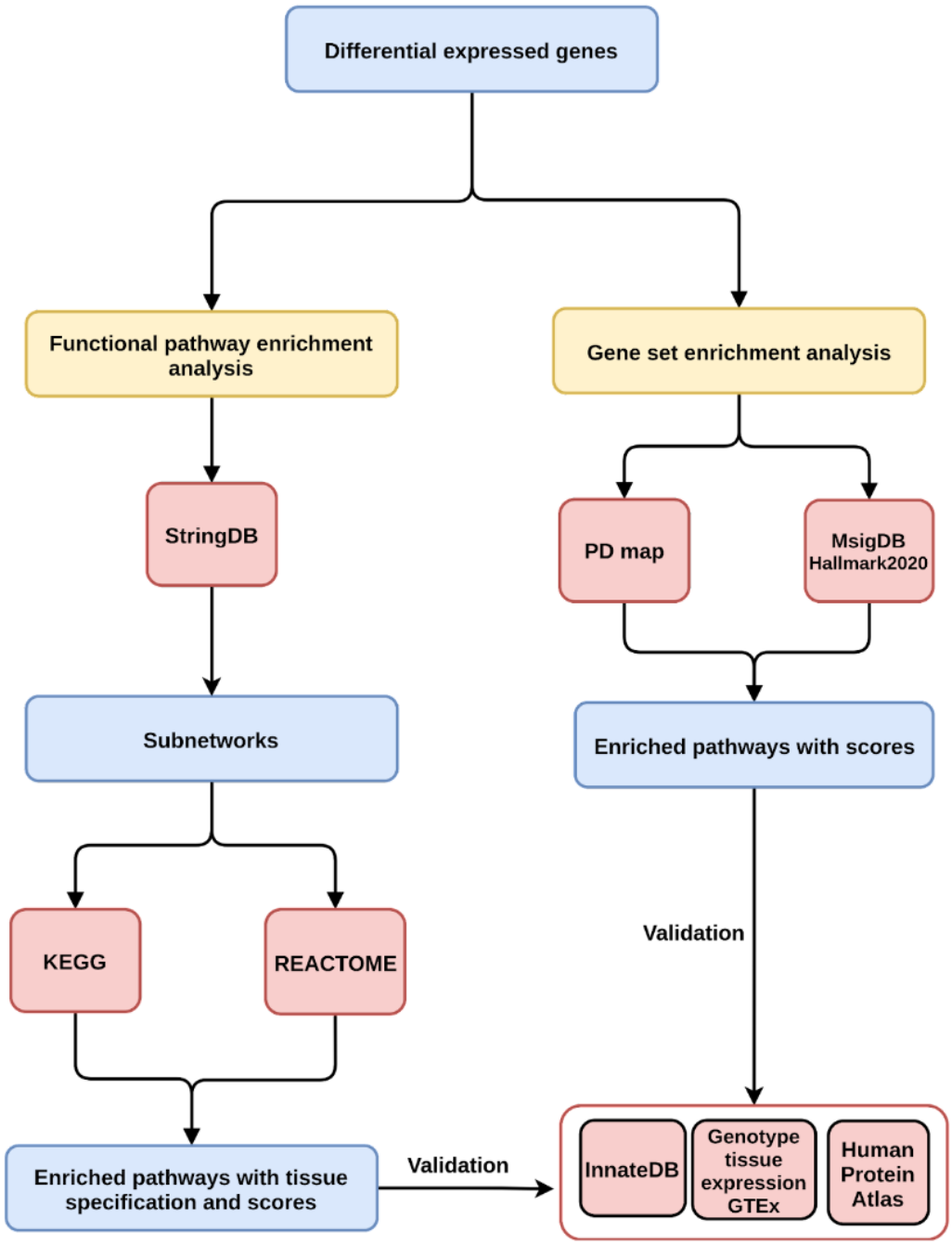


Figure B.22 Use different databases and resources for the enrichment analysis to verify the T2DM enrichment results

Pathway	Target Biomolecule	
Dopamine transcription	BDNF	knockout
	RET	knockout
	SNCA	knockout
WNT-PI3K/AKT	IGF1	knockout
	INSR	knockout
	IRS1	knockout
	MAPK1	knockout
	PHLPP1	knockout
	PTEN	knockout
	RET	knockout
	SNCA	knockout
	ROCK	overexpression
FOXO3 activity	BECN1	knockout
	BNIP3	knockout
	SIRT1	overexpression
	SNCA	knockout
mTOR pathway	MAPK1	knockout
	PHLPP1	knockout
	SIRT1	knockout
PRKN	BCL2	knockout
	BNIP3	knockout
	SNCA	knockout

Table B.3 The following table illustrates examples of perturbing T2DM-related biomolecules through knockout or overexpression on the progression of the PD cohort. The biomolecules presented in the table have been identified as being involved in enriched pathways.

Table B.4 miRNAs and parameters used in the models calibrations.

Cohort	Target.score	miRNA	gene.symbol	log2fc	abs.cohen.d	activity
SWEDD	100	hsa-miR-9-3p	SESN3	2.518635	0.213411	0.786589
SWEDD	95	hsa-miR-3121-3p	RXRA	2.306803	0.574494	0.504377
SWEDD	99	hsa-miR-9-3p	PRKAA2	2.518635	0.213411	0.786589
SWEDD	99	hsa-miR-450b-5p	PBX1	1.581469	0.13098	0.86902
SWEDD	95	hsa-miR-766-5p	GCH1	-1.80917	0.124675	0.124675
SWEDD	96	hsa-miR-582-5p	FOXO1	1.404217	0.029289	0.970711
SWEDD	99	hsa-miR-454-3p	TSC1	2.47141	0.167006	0.832994
SWEDD	98	hsa-miR-15b-5p	UBE2V1	1.491016	0.039489	0.960511
SWEDD	95	hsa-miR-96-5p	ATXN3	3.628876	0.043802	0.956198
SWEDD	97	hsa-miR-26a-5p	BAG4	3.15791	0.048562	0.951438
SWEDD	97	hsa-miR-424-5p	FBXW7	2.718901	0.111716	0.888284
SWEDD	95	hsa-miR-15b-5p	GABARAPL1	1.491016	0.039489	0.960511
SWEDD	95	hsa-miR-3121-3p	TIMM17A	2.306803	0.425506	0.574494
SWEDD	96	hsa-miR-26a-5p	ULK1	3.15791	0.048562	0.951438
SWEDD	95	hsa-miR-6515-5p	AKT1	2.810068	0.154843	0.845157
SWEDD	98	hsa-miR-20a-5p	E2F1	-3.25562	0.134403	0.134403
SWEDD	95	hsa-miR-20a-3p	EIF2AK3	-1.89509	0.014047	0.014047
SWEDD	98	hsa-miR-214-3p	GSK3B	2.59832	0.23732	0.76268
SWEDD	99	hsa-miR-130a-3p	IGF1	1.843781	0.334109	0.665891
SWEDD	96	hsa-let-7f-5p	IGF1R	1.88839	0.217186	0.782814
SWEDD	95	hsa-miR-214-3p	IRS1	2.59832	0.23732	0.76268
SWEDD	97	hsa-miR-212-3p	MAPK1	2.268402	0.057177	0.942823
SWEDD	96	hsa-miR-548t-5p	NEDD4	-2.02465	0.144947	0.144947
SWEDD	96	hsa-miR-26a-5p	PRKCD	3.15791	0.048562	0.951438
SWEDD	99	hsa-miR-26a-5p	PTEN	3.15791	0.048562	0.951438
SWEDD	95	hsa-miR-374a-5p	TFDP1	-1.99329	0.130566	0.130566
prodromal	100	hsa-miR-3121-3p	BDNF	1.633684	0.018	0.982
prodromal	99	hsa-miR-374a-5p	EN1	-2.22397	0.22	0.22
prodromal	96	hsa-miR-582-5p	FOXO1	1.139541	0.222221	0.777779
prodromal	95	hsa-miR-766-5p	GCH1	-1.74177	0.021	0.021
prodromal	99	hsa-miR-450b-5p	PBX1	0.964192	0.782969	0.217031
prodromal	99	hsa-miR-9-3p	PRKAA2	2.277933	0.851154	0.148846
prodromal	95	hsa-miR-3121-3p	RXRA	1.633684	0.981296	0.018704
prodromal	100	hsa-miR-9-3p	SESN3	2.277933	0.148846	0.851154
prodromal	95	hsa-miR-6515-5p	AKT1	2.762035	0.247214	0.752786
prodromal	99	hsa-miR-148b-3p	BCL2L11	-1.91383	0.12467	0.12467
prodromal	98	hsa-miR-374a-5p	CEBPB	-2.22397	0.225932	0.225932
prodromal	99	hsa-miR-21-5p	FASLG	-2.52095	0.120716	0.120716
prodromal	95	hsa-miR-15b-5p	GABARAPL1	1.659935	0.07363	0.889544
prodromal	95	hsa-miR-199b-3p	MAP3K5	-2.25469	0.07628	0.07628
prodromal	97	hsa-miR-214-3p	MFN2	2.586941	0.104136	0.895864
prodromal	99	hsa-miR-374b-3p	PPARGC1A	-2.9452	0.051324	0.051324
prodromal	99	hsa-miR-142-3p	RICTOR	1.116804	0.441571	0.558429
prodromal	100	hsa-miR-9-3p	SESN3	2.277933	0.148846	0.851154
prodromal	95	hsa-miR-338-5p	SIRT1	-1.75972	0.230152	0.230152
prodromal	95	hsa-miR-212-3p	SOD2	1.864045	0.113853	0.886147
prodromal	99	hsa-miR-20a-5p	TXNIP	-3.53187	0.030064	0.030064
prodromal	95	hsa-miR-6515-5p	AKT1	2.762035	0.247214	0.752786
prodromal	95	hsa-miR-96-5p	DEPTOR	3.32442	0.033239	0.966761
prodromal	98	hsa-miR-214-3p	GSK3B	2.586941	0.104136	0.895864
prodromal	97	hsa-miR-212-3p	MAPK1	1.864045	0.113853	0.886147
prodromal	97	hsa-miR-96-5p	MTOR	3.32442	0.033239	0.966761
prodromal	99	hsa-miR-548b-5p	PRKAA1	-1.47664	0.012277	0.851154
prodromal	99	hsa-miR-9-3p	PRKAA2	2.277933	0.148846	0.851154
prodromal	96	hsa-miR-338-5p	RHEB	-1.75972	0.220152	0.558429
prodromal	99	hsa-miR-142-3p	RICTOR	1.116804	0.441571	0.558429

continued on next page ...

...continued from previous page

Cohort	Target.score	miRNA	gene.symbol	log2fc	abs.cohen.d	activity
prodromal	99	hsa-miR-20a-5p	RRAGD	-3.53187	0.030064	0.876035
prodromal	95	hsa-miR-338-5p	SIRT1	-1.75972	0.220152	0.876035
prodromal	99	hsa-miR-130a-3p	TSC1	2.055169	0.123965	0.876035
prodromal	98	hsa-miR-15b-5p	UBE2V1	1.659935	0.07363	0.92637
prodromal	95	hsa-miR-96-5p	ATXN3	3.32442	0.033239	0.966761
prodromal	97	hsa-miR-26a-5p	BAG4	3.029368	0.287792	0.712208
prodromal	97	hsa-miR-424-5p	FBXW7	2.753274	0.147282	0.852718
prodromal	95	hsa-miR-15b-5p	GABARAPL1	1.659935	0.07363	0.92637
prodromal	95	hsa-miR-3121-3p	TIMM17A	1.633684	0.018704	0.981296
prodromal	96	hsa-miR-26a-5p	ULK1	3.029368	0.287792	0.712208
prodromal	95	hsa-miR-6515-5p	AKT1	2.762035	0.247214	0.752786
prodromal	98	hsa-miR-20a-5p	E2F1	-3.53187	0.030064	0.030064
prodromal	95	hsa-miR-20a-3p	EIF2AK3	-2.56318	0.213236	0.213236
prodromal	99	hsa-miR-130a-3p	IGF1	2.055169	0.123965	0.876035
prodromal	96	hsa-let-7f-5p	IGF1R	2.031796	0.048667	0.951333
prodromal	95	hsa-miR-214-3p	IRS1	2.586941	0.104136	0.895864
prodromal	97	hsa-miR-212-3p	MAPK1	1.864045	0.113853	0.886147
prodromal	96	hsa-miR-548t-5p	NEDD4	-2.8389	0.036857	0.036857
prodromal	96	hsa-miR-26a-5p	PRKCD	3.029368	0.287792	0.712208
prodromal	99	hsa-miR-26a-5p	PTEN	3.029368	0.287792	0.712208
prodromal	95	hsa-miR-374a-5p	TFDP1	-2.22397	0.225932	0.225932
parkinsonism	97	hsa-miR-1185-1-3p	SFPQ	0.945454	0.203266	0.79
parkinsonism	100	hsa-miR-181d-5p	SESN3	1.014874	0.084812	0.903731
parkinsonism	97	hsa-miR-222-3p	RGS6	0.956793	0.093351	0.913305
parkinsonism	99	hsa-miR-30e-3p	PRKAA2	0.864951	0.316977	0.86
parkinsonism	95	hsa-miR-196b-5p	PBX1	0.999252	0.001291	0.998709
parkinsonism	95	hsa-miR-485-3p	NRF1	0.871379	0.197863	0.802137
parkinsonism	95	hsa-miR-27b-3p	MAP1B	1.008988	0.015269	0.932073
parkinsonism	95	hsa-miR-377-3p	LMX1A	0.893432	0.197482	0.802518
parkinsonism	98	hsa-miR-182-5p	FOXO3	1.080294	0.138897	0.861103
parkinsonism	98	hsa-miR-486-5p	FOXO1	1.011897	0.112411	0.887589
parkinsonism	99	hsa-miR-374b-5p	EN1	1.010525	0.013122	0.983085
parkinsonism	99	hsa-miR-495-3p	BDNF	0.882131	0.207194	0.792806
parkinsonism	98	hsa-miR-485-3p	ADCYAP1	0.871379	0.197863	0.802137
parkinsonism	97	hsa-miR-30a-5p	ATG12	0.90731	0.16249	0.82592
parkinsonism	99	hsa-miR-148a-3p	BCL2L11	0.91319	0.149668	0.419232
parkinsonism	96	hsa-miR-182-5p	BNIP3	1.080294	0.138897	0.138897
parkinsonism	98	hsa-miR-374a-5p	CEBPB	1.022915	0.020708	0.033829
parkinsonism	99	hsa-miR-21-5p	FASLG	0.979987	0.038118	0.038118
parkinsonism	98	hsa-miR-182-5p	FOXO3	1.080294	0.138897	0.138897
parkinsonism	96	hsa-miR-16-5p	GABARAPL1	1.037644	0.090147	0.152009
parkinsonism	96	hsa-miR-942-5p	HSPD1	0.991871	0.019395	0.019395
parkinsonism	98	hsa-miR-139-5p	JUN	0.965532	0.081584	0.081584
parkinsonism	95	hsa-miR-199a-3p	MAP3K5	1.049396	0.055289	0.070962
parkinsonism	96	hsa-miR-766-3p	MFN2	0.982408	0.046162	0.046162
parkinsonism	100	hsa-miR-23a-3p	PPARGC1A	0.903066	0.253717	0.36427
parkinsonism	99	hsa-miR-142-3p	RICTOR	0.851202	0.226819	0.430085
parkinsonism	100	hsa-miR-23a-3p	SESN3	0.903066	0.253717	0.577613
parkinsonism	95	hsa-miR-338-5p	SIRT1	0.996662	0.014215	0.043601
parkinsonism	96	hsa-miR-377-3p	SOD2	0.893432	0.197482	0.430507
parkinsonism	99	hsa-miR-106b-5p	TXNIP	0.953409	0.122438	0.485466
parkinsonism	99	hsa-miR-671-5p	CAMKK2	1.117669	0.322512	0.677488
parkinsonism	99	hsa-miR-495-3p	DDIT4	0.882131	0.207194	0.792806
parkinsonism	100	hsa-miR-495-3p	DEPDC5	0.882131	0.207194	1
parkinsonism	95	hsa-miR-96-5p	DEPTOR	1.024287	0.033645	0.966355
parkinsonism	96	hsa-miR-26a-5p	GSK3B	0.957107	0.110998	1
parkinsonism	99	hsa-miR-543	MAPK1	0.88518	0.260174	0.739826
parkinsonism	97	hsa-miR-7-5p	MAPKAP1	0.973849	0.047077	0.952923

continued on next page ...

...continued from previous page

Cohort	Target.score	miRNA	gene.symbol	log2fc	abs.cohen.d	activity
parkinsonism	97	hsa-miR-96-5p	MTOR	1.024287	0.033645	1
parkinsonism	98	hsa-miR-7-5p	PARP1	0.973849	0.047077	0.952923
parkinsonism	97	hsa-miR-16-2-3p	PRKAA1	0.95004	0.062892	0.62013
parkinsonism	99	hsa-miR-30e-3p	PRKAA2	0.864951	0.316977	1
parkinsonism	96	hsa-miR-338-5p	RHEB	0.996662	0.014215	0.985785
parkinsonism	99	hsa-miR-142-3p	RICTOR	0.851202	0.226819	1
parkinsonism	95	hsa-miR-4677-3p	RICTOR	0.945166	0.159608	0.569915
parkinsonism	98	hsa-miR-200c-3p	RPS6KB1	1.005142	0.019552	1
parkinsonism	99	hsa-miR-17-5p	RRAGD	1.037103	0.072002	1
parkinsonism	95	hsa-miR-204-5p	SIRT1	0.989836	0.029387	1
parkinsonism	99	hsa-miR-130b-3p	TSC1	0.992711	0.027335	0.793577
parkinsonism	97	hsa-miR-497-5p	UBE2V1	1.029254	0.099932	0.591873
parkinsonism	96	hsa-miR-1271-5p	ATXN3	1.008758	0.023443	0.976557
parkinsonism	97	hsa-miR-26b-5p	BAG4	1.083644	0.103949	0.896051
parkinsonism	100	hsa-miR-32-5p	FBXW7	0.938829	0.085526	0.914474
parkinsonism	96	hsa-miR-195-5p	GABARAPL1	1.049639	0.061863	0.938137
parkinsonism	99	hsa-miR-421	TOMM70	0.985332	0.057535	0.942465
parkinsonism	96	hsa-miR-26b-5p	ULK1	1.083644	0.103949	0.896051
parkinsonism	98	hsa-miR-223-5p	VPS13C	0.985586	0.035528	0.964472
parkinsonism	98	hsa-miR-376a-3p	AGO2	0.802491	0.327577	0.672423
parkinsonism	98	hsa-miR-106b-5p	E2F1	0.953409	0.122438	0.877562
parkinsonism	95	hsa-miR-20a-3p	EIF2AK3	0.983895	0.02657	0.97343
parkinsonism	97	hsa-miR-545-3p	EIF4E	1.014214	0.04469	0.95531
parkinsonism	96	hsa-miR-26a-5p	GSK3B	0.957107	0.110998	0.889002
parkinsonism	97	hsa-miR-7-5p	IDE	0.973849	0.047077	0.952923
parkinsonism	99	hsa-miR-301a-3p	IGF1	1.01838	0.019528	0.980472
parkinsonism	97	hsa-miR-182-5p	IGF1R	1.080294	0.138897	0.861103
parkinsonism	97	hsa-miR-660-5p	IRS1	0.974744	0.046464	0.953536
parkinsonism	99	hsa-miR-543	MAPK1	0.88518	0.260174	0.739826
parkinsonism	98	hsa-miR-30a-5p	NEDD4	0.90731	0.16249	0.83751
parkinsonism	96	hsa-miR-331-3p	PHLPP1	0.942426	0.137542	0.862458
parkinsonism	97	hsa-miR-133b	PPP2CA	0.98116	0.026068	0.973932
parkinsonism	97	hsa-miR-1185-1-3p	PPP2CB	0.945454	0.203266	0.796734
parkinsonism	96	hsa-miR-26a-5p	PRKCD	0.957107	0.110998	0.889002
parkinsonism	96	hsa-miR-26b-5p	PRKCD	1.083644	0.103949	0.896051
parkinsonism	99	hsa-miR-23a-3p	PTEN	0.903066	0.253717	0.746283
parkinsonism	98	hsa-miR-30e-3p	ROCK2	0.864951	0.316977	0.683023
parkinsonism	98	hsa-miR-200c-3p	RPS6KB1	1.005142	0.019552	0.980448
parkinsonism	95	hsa-miR-374a-5p	TFDP1	1.022915	0.020708	0.979292